PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A61K 38/08, 38/10, 39/10, 39/02, 39/12, C07K 7/00, 14/005, 14/20, 14/195, 14/725, C07H 21/04

A1

(11) International Publication Number: WO 98/32456

(43) International Publication Date: 30 July 1998 (30.07.98)

(21) International Application Number: PCT/US98/01373

(22) International Filing Date: 23 January 1998 (23.01.98)

(30) Priority Data:

60/036,713	23 January 1997 (23.01.97)	US
60/037,432	7 February 1997 (07.02.97)	US

(71) Applicant *(for all designated States except US)*: EPIMMUNE, INC. [US/US]; Suite 200, 6555 Nancy Ridge Drive, San Diego, CA 92121 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: SETTE, Alessandro [IT/US]; 5551 Linda Rosa Avenue, La Jolla, CA 92037 (US). SIDNEY, John [US/US]; 8541-D Villa La Jolla Drive, La Jolla, CA 92037 (US). SOUTHWOOD, Scott [US/US]; 10679 Strathmore Drive, Santee, CA 92071 (US).

(74) Agent: BASTIAN, Kevin, L.; Townsend and Townsend and Crew LLP, 8th floor, Two Embarcadero Center, San Francisco, CA 94111 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: IDENTIFICATION OF BROADLY REACTIVE DR RESTRICTED EPITOPES

(57) Abstract

The present invention is based on peptide binding specificities of HLA DR4w4, DR1 and DR7. Peptides binding to these DR molecules have a motif characterized by a large aromatic or hydrophobic residue in position 1 (Y, F, W, L, I, V, M) and a small, non charged residue in position 6 (S, T, C, A, P, V, I, L, M). In addition, allele–specific secondary effects and secondary anchors are defined, and these results were utilized to derive allele specific algorithms. By the combined use of such alogirthms peptides capable of degenerate DR1, 4, 7 binding were identified.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

IDENTIFICATION OF BROADLY REACTIVE DR RESTRICTED EPITOPES

CROSS-REFERENCE TO RELATED APPLICATIONS

The present application is a continuation in part of USSN 60/036,713, filed January 23, 1997 and 60/037,432 filed February 7, 1997, both of which are incorporated herein by reference.

BACKGROUND OF THE INVENTION

Helper T lymphocytes (HTL) play several important functions in immunity to pathogens. Firstly, they provide help for induction of both CTL and antibody responses. By both direct contact and by secreting lymphokines such as IL2 and IL4, HTL promote and support the expansion and differentiation of T and B cell precursors into effector cells. In addition, HTL can also be effectors in their own right, an activity also mediated by direct cell contact and secretion of lymphokines, such as IFNγ and TNFα. HTL have been shown to have direct effector activity in case of tumors, as well as viral, bacterial, parasitic, and fungal infections.

HTL recognize a complex formed between Class II MHC molecules and antigenic peptides, usually between 10 and 20 residues long, and with an average size of between 13 and 16 amino acids. Peptide-Class II interactions have been analyzed in detail, both at the structural and functional level, and peptide motifs specific for various human and mouse Class II molecules have been proposed.

In the last few years, epitope based vaccines have received considerable attention as a possible mean to develop novel prophylactic vaccines and immunotherapeutic strategies. Selection of appropriate T and B cell epitopes should allow to focus the immune system toward conserved epitopes of pathogens which are characterized by high sequence variability (such as HIV, HCV and Malaria).

In addition, focusing the immune response towards selected determinants could be of value in the case of various chronic viral diseases and cancer, where T cells directed against the immunodominant epitopes might have been inactivated while T cells specific for subdominant epitopes might have escaped T cell tolerance. The use of epitope

based vaccines also allows to avoid "suppressive" T cell determinants which induce TH_2 responses, in conditions where a TH_1 response is desirable, or vice versa.

Finally, epitope based vaccines also offer the opportunity to include in the vaccine construct epitopes that have been engineered to modulate their potency, either by increasing MHC binding affinity, or by alteration of its TCR contact residues, or both. Inclusion of completely synthetic non-natural or generically unrelated to the pathogen epitopes (such as TT derived "universal" epitopes), also represents a possible mean of modulating the HTL response toward a TH_1, or TH_2 phenotype.

Once appropriate epitope determinants have been defined, they can be assorted and delivered by various means, which include lipopeptides, viral delivery vectors, particles of viral or synthetic origin, naked or particle absorbed cDNA.

However, before appropriate epitopes can be defined, one major obstacle has to be overcome, namely the very high degree of polymorphism of the MHC molecules expressed in the human population. In fact, more than two hundred different types of HLA Class I and Class II molecules have already been identified. It has been demonstrated that in the case of HLA Class I molecules, peptides capable of binding several different HLA Class I molecules can be identified. Over 60% of the known HLA Class I molecules can, in fact, be grouped in four broad HLA supertypes, characterized by similar peptide binding specificities (HLA supermotifs).

In the case of Class III molecules, it is also known that peptides capable of binding multiple HLA types and of being immunogenic in the context of different HLA molecules do indeed exist. Until now, however, a general method for their identification has not been developed, probably at least in part a reflection of the fact that quantitative DR binding assays are labor intensive and that a large number of alleles must to be considered.

The present invention addresses these and other needs.

SUMMARY OF THE INVENTION

The present invention is based, at least in part, on the discovery and validation of specific motifs and assay systems for various DR molecules, representative of the worldwide population. Their application to the identification of broadly degenerate HLA Class II binding peptides is also described.

Definitions

The term "peptide" is used interchangeably with "oligopeptide" in the present specification to designate a series of residues, typically L-amino acids, connected one to the other typically by peptide bonds between the alpha-amino and carbonyl groups of adjacent amino acids. The oligopeptides of the invention are less than about 50 residues in length and usually consist of between about 10 and about 30 residues, more usually between about 12 and 25, and often between about 15 and about 20 residues.

An "immunogenic peptide" is a peptide which comprises an allele-specific motif such that the peptide will bind an MHC molecule and induce a HTL response. Immunogenic peptides of the invention are capable of binding to an appropriate HLA molecule and inducing HTL response against the antigen from which the immunogenic peptide is derived.

A "conserved residue" is a conserved amino acid occupying a particular position in a peptide motif typically one where the MHC structure may provide a contact point with the immunogenic peptide. One to three, typically two, conserved residues within a peptide of defined length defines a motif for an immunogenic peptide. These residues are typically in close contact with the peptide binding groove, with their side chains buried in specific pockets of the groove itself.

The term "motif" refers to the pattern of residues of defined length, usually between about 8 to about 11 amino acids, which is recognized by a particular MHC allele.

The term "supermotif" refers to motifs that, when present in an immunogenic peptide, allow the peptide to bind more than one HLA antigen. The supermotif preferably is recognized by at least one HLA allele having a wide distribution in the human population, preferably recognized by at least two alleles, more preferably recognized by at least three alleles, and most preferably recognized by more than three alleles.

The phrases "isolated" or "biologically pure" refer to material which is substantially or essentially free from components which normally accompany it as found in its native state. Thus, the peptides of this invention do not contain materials normally associated with their *in situ* environment, e.g., MHC I molecules on antigen presenting cells. Even where a protein has been isolated to a homogenous or dominant band, there are trace contaminants in the range of 5-10% of native protein which co-purify with the desired protein. Isolated peptides of this invention do not contain such endogenous co-purified protein.

The term "residue" refers to an amino acid or amino acid mimetic incorporated in an oligopeptide by an amide bond or amide bond mimetic.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 presents a map of the positive or negative effect of each of the 20 naturally occurring amino acids on DR4w4 binding capacity when occupying a particular position, relative to the main P1-P6 anchors.

Figure 2A presents a map of the positive or negative effect of each of the 20 naturally occurring amino acids on DR1 binding capacity when occupying a particular position, relative to the main P1-P6 anchors.

Figure 2B presents a map of the positive or negative effect of each of the 20 naturally occurring amino acids on DR7 binding capacity when occupying a particular position, relative to the main P1-P6 anchors.

DESCRIPTION OF THE PREFERRED EMBODIMENT

The present invention relates to compositions and methods for preventing, treating or diagnosing a number of pathological states such as viral, fungal, bacterial and parasitic diseases and cancers. In particular, it provides novel peptides capable of binding selected major histocompatibility complex (MHC) class II molecules and inducing an immune response.

Peptide binding to MHC molecules is determined by the allelic type of the MHC molecule and the amino acid sequence of the peptide. MHC class I-binding peptides usually contain within their sequence two conserved ("anchor") residues that interact with corresponding binding pockets in the MHC molecule. Specific combination of anchor

residues (usually referred to as "MHC motifs") required for binding by several allelic forms of human MHC (HLA, histocompatibility leukocyte antigens) are described in International Applications WO 94/03205 and WO 94/20127. Definition of specific MHC motifs allows one to predict from the amino acid sequence of an individual protein, which peptides have the potential of being immunogenic for CTL. These applications describe methods for preparation and use of immunogenic peptides in the treatment of disease. The peptides described here can also be used as helper T peptides in combination with peptides which induce a CTL response. This is described in WO 95/07077.

The DR-binding peptides of the present invention or nucleic acids encoding them can be administered to mammals, particularly humans, for prophylactic and/or therapeutic purposes. The DR peptides can be used to enhance immune responses against other immunogens administered with the peptides, when the peptides of the invention are used as helper peptides. For instance, mixtures of peptides of the invention in combination with peptides that induce CTL responses may be used to treat and/or prevent viral infection and cancer. Alternatively, immunogens which induce antibody responses can be used. Examples of diseases which can be treated using the immunogenic mixtures of DR peptides and other immunogens include prostate cancer, hepatitis B, hepatitis C, AIDS, renal carcinoma, cervical carcinoma, lymphoma, CMV and condyloma acuminatum.

The DR-binding peptides or nucleic acids encoding them may also be used to treat a variety of conditions involving unwanted T cell reactivity. Examples of diseases which can be treated using DR-binding peptides include autoimmune diseases (e.g., rheumatoid arthritis, multiple sclerosis, and myasthenia gravis), allograft rejection, allergies (e.g., pollen allergies), lyme disease, hepatitis, LCMV, post-streptococcal endocarditis, or glomerulonephritis, and food hypersensitivities.

In therapeutic applications, the immunogenic compositions or the DR-binding peptides or nucleic acids of the invention are administered to an individual already suffering from cancer, autoimmune disease, or infected with the virus of interest. Those in the incubation phase or the acute phase of the disease may be treated with the DR-binding peptides or immunogenic conjugates separately or in conjunction with other treatments, as appropriate.

In therapeutic applications, compositions comprising immunogenic compositions are administered to a patient in an amount sufficient to elicit an effective immune response to the virus or tumor antigen and to cure or at least partially arrest symptoms and/or complications. Similarly, compositions comprising DR-binding peptides are administered in an amount sufficient to cure or at least partially arrest the symptoms of the disease and its complications. An amount adequate to accomplish this is defined as "therapeutically effective dose." Amounts effective for this use will depend on, e.g., the peptide composition, the manner of administration, the stage and severity of the disease being treated, the weight and general state of health of the patient, and the judgment of the prescribing physician.

Therapeutically effective amounts of the immunogenic compositions of the present invention generally range for the initial immunization (that is for therapeutic or prophylactic administration) from about 1.0 μg to about 10,000 μg of peptide for a 70 kg patient, usually from about 100 to about 8000 μg, and preferably between about 200 and about 6000 μg. These doses are followed by boosting dosages of from about 1.0 μg to about 1000 μg of peptide pursuant to a boosting regimen over weeks to months depending upon the patient's response and condition by measuring specific immunogenic activity in the patient's blood.

It must be kept in mind that the compositions of the present invention may generally be employed in serious disease states, that is, life-threatening or potentially life-threatening situations. In such cases, in view of the minimization of extraneous substances and the relative nontoxic nature of the conjugates, it is possible and may be felt desirable by the treating physician to administer substantial excesses of these compositions.

For prophylactic use, administration should be given to risk groups. For example, protection against malaria, hepatitis, or AIDS may be accomplished by prophylactically administering compositions of the invention, thereby increasing immune capacity. Therapeutic administration may begin at the first sign of disease or the detection or surgical removal of tumors or shortly after diagnosis in the case of acute infection. This is followed by boosting doses until at least symptoms are substantially abated and for a period thereafter. In chronic infection, loading doses followed by boosting doses may be required.

Treatment of an infected individual with the compositions of the invention may hasten resolution of the infection in acutely infected individuals. For those individuals susceptible (or predisposed) to developing chronic infection the compositions are particularly useful in methods for preventing the evolution from acute to chronic infection. Where the susceptible individuals are identified prior to or during infection, for instance, as described herein, the composition can be targeted to them, minimizing need for administration to a larger population.

The peptide mixtures or conjugates can also be used for the treatment of chronic infection and to stimulate the immune system to eliminate virus-infected cells in carriers. It is important to provide an amount of immuno-potentiating peptide in a formulation and mode of administration sufficient to effectively stimulate a cytotoxic T cell response. Thus, for treatment of chronic infection, a representative dose is in the range of about 1.0 μg to about 5000 μg, preferably about 5 μg to 1000 μg for a 70 kg patient per dose. Immunizing doses followed by boosting doses at established intervals, e.g., from one to four weeks, may be required, possibly for a prolonged period of time to effectively immunize an individual. In the case of chronic infection, administration should continue until at least clinical symptoms or laboratory tests indicate that the viral infection has been eliminated or substantially abated and for a period thereafter.

The pharmaceutical compositions for therapeutic or prophylactic treatment are intended for parenteral, topical, oral or local administration. Typically, the pharmaceutical compositions are administered parenterally, e.g., intravenously, subcutaneously, intradermally, or intramuscularly. Because of the ease of administration, the vaccine compositions of the invention are particularly suitable for oral administration. Thus, the invention provides compositions for parenteral administration which comprise a solution of the peptides or conjugates dissolved or suspended in an acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers may be used, e.g., water, buffered water, 0.9% saline, 0.3% glycine, hyaluronic acid and the like. These compositions may be sterilized by conventional, well known sterilization techniques, or may be sterile filtered. The resulting aqueous solutions may be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile solution prior to administration. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions, such as pH adjusting and

buffering agents, tonicity adjusting agents, wetting agents and the like, for example, sodium acetate, sodium lactate, sodium chloride, potassium chloride, calcium chloride, sorbitan monolaurate, triethanolamine oleate, etc.

The concentration of DR and/or CTL stimulatory peptides of the invention in the pharmaceutical formulations can vary widely, i.e., from less than about 0.1%, usually at or at least about 2% to as much as 20% to 50% or more by weight, and will be selected primarily by fluid volumes, viscosities, etc., in accordance with the particular mode of administration selected.

The peptides and conjugates of the invention may also be administered via liposomes, which serve to target the conjugates to a particular tissue, such as lymphoid tissue, or targeted selectively to infected cells, as well as increase the half-life of the peptide composition. Liposomes include emulsions, foams, micelles, insoluble monolayers, liquid crystals, phospholipid dispersions, lamellar layers and the like. In these preparations the peptide to be delivered is incorporated as part of a liposome, alone or in conjunction with a molecule which binds to, e.g., a receptor prevalent among lymphoid cells, such as monoclonal antibodies which bind to the CD45 antigen, or with other therapeutic or immunogenic compositions. Thus, liposomes filled with a desired peptide or conjugate of the invention can be directed to the site of lymphoid cells, where the liposomes then deliver the selected therapeutic/immunogenic peptide compositions. Liposomes for use in the invention are formed from standard vesicle-forming lipids, which generally include neutral and negatively charged phospholipids and a sterol, such as cholesterol. The selection of lipids is generally guided by consideration of, e.g., liposome size, acid lability and stability of the liposomes in the blood stream. A variety of methods are available for preparing liposomes, as described in, e.g., Szoka, *et al., Ann. Rev. Biophys. Bioeng*. 9, 467 (1980), U.S. Patent Nos. 4,235,871, 4,501,728, 4,837,028, and 5,019,369, incorporated herein by reference.

For targeting to the immune cells, a ligand to be incorporated into the liposome can include, e.g., antibodies or fragments thereof specific for cell surface determinants of the desired immune system cells. A liposome suspension containing a peptide or conjugate may be administered intravenously, locally, topically, etc. in a dose which varies according to, inter alia, the manner of administration, the conjugate being delivered, and the stage of the disease being treated.

Alternatively, DNA or RNA encoding one or more DR peptides and a polypeptide containing one or more CTL epitopes or antibody inducing epitopes may be introduced into patients to obtain an immune response to the polypeptides which the nucleic acid encodes. Wolff, *et. al., Science* 247: 1465-1468 (1990) describes the use of nucleic acids to produce expression of the genes which the nucleic acids encode. Such use is also disclosed in U.S. Patent Nos. 5,580,859 and 5,589,466. The nucleic acids can also be administered using ballistic delivery as described, for instance, in U.S. Patent No. 5,204,253. Particles comprised solely of DNA can be administered. Alternatively, DNA can be adhered to particles, such as gold particles. The nucleci acids can also be delivered complexed to cationic compounds, such as cationic lipids. Lipid-mediated gene delivery methods are described, for instance, in WO 96/18372; WO 93/24640; Mannino and Gould-Fogerite (1988) *BioTechniques* 6(7): 682-691; Rose U.S. Pat No. 5,279,833; WO 91/06309; and Felgner *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84: 7413-7414. The peptides of the invention can also be expressed by attenuated viral hosts, such as vaccinia or fowlpox. This approach involves the use of vaccinia virus as a vector to express nucleotide sequences that encode the peptides of the invention. Upon introduction into an acutely or chronically infected host or into a noninfected host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits a host CTL response. Vaccinia vectors and methods useful in immunization protocols are described in, e.g.,U.S. Patent No. 4,722,848, incorporated herein by reference. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover et al. (Nature 351:456-460 (1991)) which is incorporated herein by reference. A wide variety of other vectors useful for therapeutic administration or immunization of the peptides of the invention, e.g., Salmonella typhi vectors and the like, will be apparent to those skilled in the art from the description herein.

A preferred means of administering nucleic acids encoding the peptides of the invention uses minigene constructs encoding multiple peptides of the invention along with CTL inducing peptides. To create a DNA sequence encoding the selected DR peptides and CTL epitopes for expression in human cells, the amino acid sequences of the epitopes are reverse translated. A human codon usage table is used to guide the codon choice for each amino acid. These epitope-encoding DNA sequences are directly adjoined, creating a continuous polypeptide sequence. To optimize expression and/or

immunogenicity, additional elements can be incorporated into the minigene design. Examples of amino acid sequence that could be reverse translated and included in the minigene sequence include: DR peptides of the invention, a leader (signal) sequence, one or more CTL epitope, and an endoplasmic reticulum retention signal. In addition, MHC presentation of CTL epitopes may be improved by including synthetic (e.g. poly-alanine) or naturally-occurring flanking sequences adjacent to the CTL epitopes.

The minigene sequence is converted to DNA by assembling oligonucleotides that encode the plus and minus strands of the minigene. Overlapping oligonucleotides (30-100 bases long) are synthesized, phosphorylated, purified and annealed under appropriate conditions using well known techniques. he ends of the oligonucleotides are joined using T4 DNA ligase. This synthetic minigene, encoding the CTL epitope polypeptide, can then cloned into a desired expression vector.

Standard regulatory sequences well known to those of skill in the art are included in the vector to ensure expression in the target cells. Several vector elements are required: a promoter with a down-stream cloning site for minigene insertion; a polyadenylation signal for efficient transcription termination; an *E. coli* origin of replication; and an *E. coli* selectable marker (e.g. ampicillin or kanamycin resistance). Numerous promoters can be used for this purpose, *e.g.,* the human cytomegalovirus (hCMV) promoter. *See*, U.S. Patent Nos. 5,580,859 and 5,589,466 for other suitable promoter sequences.

Additional vector modifications may be desired to optimize minigene expression and immunogenicity. In some cases, introns are required for efficient gene expression, and one or more synthetic or naturally-occurring introns could be incorporated into the transcribed region of the minigene. The inclusion of mRNA stabilization sequences can also be considered for increasing minigene expression. It has recently been proposed that immunostimulatory sequences (ISSs or CpGs) play a role in the immunogenicity of DNA vaccines. These sequences could be included in the vector, outside the minigene coding sequence, if found to enhance immunogenicity.

In some embodiments, a bicistronic expression vector, to allow production of the minigene-encoded epitopes and a second protein included to enhance or decrease immunogenicity can be used. Examples of proteins or polypeptides that could beneficially enhance the immune response if co-expressed include cytokines (e.g., IL2, IL12, GM-

CSF), cytokine-inducing molecules (e.g. LeIF) or costimulatory molecules. The HTL epitopes of the invention could be joined to intracellular targeting signals and expressed separately from the CTL epitopes. This would allow direction of the HTL epitopes to a cell compartment different than the CTL epitopes. If required, this could facilitate more efficient entry of HTL epitopes into the MHC class II pathway, thereby improving CTL induction. In contrast to CTL induction, specifically decreasing the immune response by co-expression of immunosuppressive molecules (e.g. TGF-β) may be beneficial in certain diseases.

Once an expression vector is selected, the minigene is cloned into the polylinker region downstream of the promoter. This plasmid is transformed into an appropriate *E. coli* strain, and DNA is prepared using standard techniques. The orientation and DNA sequence of the minigene, as well as all other elements included in the vector, are confirmed using restriction mapping and DNA sequence analysis. Bacterial cells harboring the correct plasmid can be stored as a master cell bank and a working cell bank.

Therapeutic quantities of plasmid DNA are produced by fermentation in *E. coli*, followed by purification. Aliquots from the working cell bank are used to inoculate fermentation medium (such as Terrific Broth), and grown to saturation in shaker flasks or a bioreactor according to well known techniques. Plasmid DNA can be purified using standard bioseparation technologies such as solid phase anion-exchange resins supplied by Quiagen. If required, supercoiled DNA can be isolated from the open circular and linear forms using gel electrophoresis or other methods.

Purified plasmid DNA can be prepared for injection using a variety of formulations. The simplest of these is reconstitution of lyophilized DNA in sterile phosphate-buffer saline (PBS). A variety of methods have been described, and new techniques may become available. As noted above, nucleic acids are conveniently formulated with cationic lipids. In addition, glycolipids, fusogenic liposomes, peptides and compounds referred to collectively as protective, interactive, non-condensing (PINC) could also be complexed to purified plasmid DNA to influence variables such as stability, intramuscular dispersion, or trafficking to specific organs or cell types.

Target cell sensitization can be used as a functional assay for expression and MHC class I presentation of minigene-encoded CTL epitopes. The plasmid DNA is

introduced into a mammalian cell line that is suitable as a target for standard CTL chromium release assays. The transfection method used will be dependent on the final formulation. Electroporation can be used for "naked" DNA, whereas cationic lipids allow direct *in vitro* transfection. A plasmid expressing green fluorescent protein (GFP) can be co-transfected to allow enrichment of transfected cells using fluorescence activated cell sorting (FACS). These cells are then chromium-51 labeled and used as target cells for epitope-specific CTL lines. Cytolysis, detected by 51Cr release, indicates production of MHC presentation of minigene-encoded CTL epitopes.

In vivo immunogenicity is a second approach for functional testing of minigene DNA formulations. Transgenic mice expressing appropriate human MHC molecules are immunized with the DNA product. The dose and route of administration are formulation dependent (e.g. IM for DNA in PBS, IP for lipid-complexed DNA). Twenty-one days after immunization, splenocytes are harvested and restimulated for 1 week in the presence of peptides encoding each epitope being tested. These effector cells (CTLs) are assayed for cytolysis of peptide-loaded, chromium-51 labeled target cells using standard techniques. Lysis of target cells sensitized by MHC loading of peptides corresponding to minigene-encoded epitopes demonstrates DNA vaccine function for *in vivo* induction of CTLs.

For solid compositions, conventional nontoxic solid carriers may be used which include, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharin, talcum, cellulose, glucose, sucrose, magnesium carbonate, and the like. For oral administration, a pharmaceutically acceptable nontoxic composition is formed by incorporating any of the normally employed excipients, such as those carriers previously listed, and generally 10-95% of active ingredient, that is, one or more conjugates of the invention, and more preferably at a concentration of 25%-75%.

For aerosol administration, the peptides are preferably supplied in finely divided form along with a surfactant and propellant. Typical percentages of conjugates are 0.01%-20% by weight, preferably 1%-10%. The surfactant must, of course, be nontoxic, and preferably soluble in the propellant. Representative of such agents are the esters or partial esters of fatty acids containing from 6 to 22 carbon atoms, such as caproic, octanoic, lauric, palmitic, stearic, linoleic, linolenic, olesteric and oleic acids with an aliphatic polyhydric alcohol or its cyclic anhydride. Mixed esters, such as mixed or

natural glycerides may be employed. The surfactant may constitute 0.1%-20% by weight of the composition, preferably 0.25-5%. The balance of the composition is ordinarily propellant. A carrier can also be included, as desired, as with, e.g., lecithin for intranasal delivery.

In another aspect the present invention is directed to vaccines which contain as an active ingredient an immunogenically effective amount of an immunogenic DR peptide or a CTL\DR peptide conjugate or nucleic acid encoding them as described herein. The conjugate(s) may be introduced into a host, including humans, linked to its own carrier or as a homopolymer or heteropolymer of active peptide units. Such a polymer has the advantage of increased immunological reaction and, where different peptides are used to make up the polymer, the additional ability to induce antibodies and/or CTLs that react with different antigenic determinants of the virus or tumor cells. Useful carriers are well known in the art, and include, e.g., thyroglobulin, albumins such as bovine serum albumin, tetanus toxoid, polyamino acids such as poly(lysine:glutamic acid), hepatitis B virus core protein, hepatitis B virus recombinant vaccine and the like. The vaccines can also contain a physiologically tolerable (acceptable) diluent such as water, phosphate buffered saline, or saline, and further typically include an adjuvant. Adjuvants such as incomplete Freund's adjuvant, aluminum phosphate, aluminum hydroxide, or alum are materials well known in the art. And, as mentioned above, CTL responses can be primed by conjugating peptides of the invention to lipids, such as P_3CSS. Upon immunization with a peptide composition as described herein, via injection, aerosol, oral, transdermal or other route, the immune system of the host responds to the vaccine by producing large amounts of CTLs specific for the desired antigen, and the host becomes at least partially immune to later infection, or resistant to developing chronic infection.

Vaccine compositions containing the DR peptides of the invention are administered to a patient susceptible to or otherwise at risk of disease, such as viral infection or cancer to elicit an immune response against the antigen and thus enhance the patient's own immune response capabilities, for instance with CTL epitopes described in **. Such an amount is defined to be an "immunogenically effective dose." In this use, the precise amounts again depend on the patient's state of health and weight, the mode of administration, the nature of the formulation, etc., but generally range from about 1.0 μg

to about 5000 μg per 70 kilogram patient, more commonly from about 10 μg to about 500 μg per 70 kg of body weight.

In some instances it may be desirable to combine the peptide vaccines of the invention with vaccines which induce neutralizing antibody responses to the virus of interest, particularly to viral envelope antigens. For instance, PADRE peptides can be combined with hepatitis vaccines to increase potency or broaden population coverage. Suitable hepatitis vaccines that can be used in this manner include, Recombivax HB® (Merck) and Engerix-B (Smith-Kline).

For therapeutic or immunization purposes, the peptides of the invention can also be expressed by attenuated viral hosts, such as vaccinia or fowlpox. This approach involves the use of vaccinia virus as a vector to express nucleotide sequences that encode the peptides of the invention. Upon introduction into an acutely or chronically infected host or into a non-infected host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits a host CTL response. Vaccinia vectors and methods useful in immunization protocols are described in, e.g., U.S. Patent No. 4,722,848, incorporated herein by reference. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover *et al., Nature* 351, 456-460 (1991)) which is incorporated herein by reference. A wide variety of other vectors useful for therapeutic administration or immunization of the peptides of the invention, e.g., Salmonella typhi vectors and the like, will be apparent to those skilled in the art from the description herein.

Antigenic conjugates may be used to elicit CTL *ex vivo*, as well. The resulting CTL can be used to treat chronic infections (viral or bacterial) or tumors in patients that do not respond to other conventional forms of therapy, or will not respond to a peptide vaccine approach of therapy. *Ex vivo* CTL responses to a particular pathogen (infectious agent or tumor antigen) are induced by incubating in tissue culture the patient's CTL precursor cells (CTLp) together with a source of antigen-presenting cells (APC) and the appropriate immunogenic peptide. After an appropriate incubation time (typically 1-4 weeks), in which the CTLp are activated and mature and expand into effector CTL, the cells are infused back into the patient, where they will destroy their specific target cell (an infected cell or a tumor cell).

The peptides of this invention may also be used to make monoclonal antibodies. Such antibodies may be useful as potential diagnostic or therapeutic agents.

The peptides may also find use as diagnostic reagents. For example, a peptide of the invention may be used to determine the susceptibility of a particular individual to a treatment regimen which employs the peptide or related peptides, and thus may be helpful in modifying an existing treatment protocol or in determining a prognosis for an affected individual. In addition, the peptides may also be used to predict which individuals will be at substantial risk for developing chronic infection.

Examples

Materials and Methods

Cells. The following Epstein-Barr virus (EBV) transformed homozygous cell lines were used as sources of human HLA Class II molecules: LG2 [DRB1c0101 (DR1)1; GM3107 [DRB50101 (DR2w2a)]; MAT (DRB10301 (DR3)1; PREISS [DRB10401 (DR4w4)1; BIN40 [DRB10404 (DR4w14)1; SWEIG [DRB11101 (DR5w11)]; PITOUT [DRB10701 (DR7)] (a); KT3 [DRB10405 (DR4w15)]; Herluf [DRB11201 (DR5w12)]; HO301 [DRB11302 (DR6w19)]; OLL [DRB10802 (DR8w2)]; and HTC9074 [DRB10901 (DR9), supplied as a kind gift by Dr. Paul Harris, Columbia University]. In some instances, transfected fibroblasts were used: L466.1 [DRB11501 (DR2w2b)]; TR81.19 [DRB30101 (DR52a)]; and L257.6 [DRB40101 (DRw53)]. (Valli, *et al. J. Clin. Invest.* 91:616 (1993). Cells were maintained *in vitro* by culture in RPMI 1640 medium supplemented with 2mM L-glutamine [GIBCO, Grand Island, NY], 50μM 2-ME, and 10% heat-inactivated FCS [Irvine Scientific, Santa Ana, CA]. Cells were also supplemented with 100 μg/ml of streptomycin and 10OU/ml of penicillin [Irvine Scientific]. Large quantities of cells were grown in spinner cultures.

Cells were lysed at a concentration of 10^8 cells/ml in PBS containing 1% NP-40 [Fluka Biochemika, Buchs, Switzerland], 1mM PMSF [CalBioChem, La Jolla, CA], 5mM Na-orthovanadate, and 25mM iodoacetamide [Sigma Chemical, St. Louis, Mo]. The lysates were cleared of debris and nuclei by centrifugation at 10,000 x *g* for 20 min.

Affinity purification of HLA-DR molecules. Class II molecules were purified by affinity chromatography as previously described (Sette, *et al. J. Immunol.* 142:35 (1989) and Gorga, *et al. J. Biol. Chem.* 262:16087 (1987)) using the mAb LB3.1 coupled to Sepharose 4B beads. Lysates were filtered through 0.8 and 0.4 μM filters and then passed

over the anti-DR column, which were then washed with 15-column volumes of 10mM TRIS in 1% NP-40, PBS and 2-column volumes of PBS containing 0.4% n-octylglucoside. Finally, the DR was eluted with 5OmM diethylamine in 0.15M NaCl containing 0.4% n-octylglucoside, pH 11.5. A 1/25 volume of 2.0M Tris, pH 6.8, was added to the eluate to reduce the pH to ~8.0, and then concentrated by centrifugation in Centriprep 30 concentrators at 2000 rpm (Amicon, Beverly, MA).

Class II peptide-binding assays. A panel of 13 different specific DR-peptide assays were utilized in the present study. These assays were chosen as to be representative of the most common DR alleles. Table I lists for each DR antigen, the representative allelic product utilized, the cell line utilized as a source of DR, and the radiolabled probe utilized in the assay. Purified human Class II molecules [5 to 500 nM] were incubated with various unlabeled peptide inhibitors and 1-10 nM ^{125}I-radiolabeled probe peptides for 48h in PBS containing 5% DMSO in the presence of a protease inhibitor cocktail. The radiolabeled probes used were HA Y307-319 (DR1), Tetanus Toxoid[TT] 830-843 (DR2w2a, DR5w111, DR7, DR8w2, DR8w3, DR9), MBP Y85-100 (DR2w2b), TT1272-1284 (DR52a), MT 65 kD Y3-13 with Y7 substituted with F for DR3, a non-natural peptide with the sequence YARFQSQTTLKQKT (DR4w4, DR4w15, DRw53) (Valli, *et al. supra*), and for DR5w12, a naturally processed peptide eluted from the cell line C1R, EALIHQLINPYVLS (DR5w12) and 650.22 peptide, (TT 830-843 A → S836 analog), for DR6w19.

Radiolabeled peptides were iodinated using the chloramine-T method. Peptide inhibitors were typically tested at concentrations ranging from 120l μg/ml to 1.2 ng/ml. The data were then plotted and the dose yielding 50% inhibition (IC50) was measured. In appropriate stoichiometric conditions, the IC50 of an unlabeled test peptide to the purified DR is a reasonable approximation of the affinity of interaction (Kd). Peptides were tested in two to four completely independent experiments. The final concentrations of protease inhibitors were: 1mM PMSF, 1.3nM 1.10 phenanthroline, 73 μM pepstatin A, 8mM EDTA, and 200 μM N alpha-p-tosyl-L-lysine chloromethyl ketone (TLCK) [All protease inhibitors from CalBioChem, La Jolla, CA]. Final detergent concentration in the incubation mixture was 0.05% Nonidet P-40. Assays were performed at pH 7.0 with the exception of DR3, which was performed at pH 4.5, and DRw53, which

was performed at pH 5.0. The pH was adjusted as previously described (Sette, *et al. J. Immunol.* 148:844 (1992)).

Class II peptide complexes were separated from free peptide by gel filtration on TSK2000 columns (TosoHaas 16215, Montgomeryville, PA), and the fraction of bound peptide calculated as previously described (Sette, *et al.*, (1989) *supra*). In preliminary experiments, the DR prep was titered in the presence of fixed amounts of radiolabeled peptides to determine the concentration of Class II molecules necessary to bind 10-20% of the total radioactivity. All subsequent inhibition and direct binding assays were the performed using these Class II concentrations.

DRB1 specificity of DR4w15, DR6w19, DR8w2, DR8w3, and DR9 assays.

Because the antibody used for purification is α-chain specific, β1 molecules are not separated from β3 (and/or β4 and β5) molecules. Development and validation of assays in regard with DRβ chain specificity has been described in detail elsewhere for many of the DR alleles listed above (108). Herein we describe for the first time DR4w15, DR6w19, DR8w2, DR8w3, and DR9 assays. Experiments addressing the β chain specificity of these new assays are described in the present section.

DR4w15. The β4 product DRw53 is co-expressed with DR4w15 and the determination of the specificity of the DR4w15 binding assay is complicated in that the same radiolabeled ligand is used for both the DR4w15 and DRw53 binding assays. Since typically β1 chains are expressed at 5-10 fold higher levels than other β chains, and all binding assays are performed utilizing limiting DR amounts, it would be predicted that the dominant specificity detected in the assay would be DR4w15. To verify that this was indeed the case, the binding pattern of a panel of 58 different synthetic peptides in the putative DR4w15 specific assay with that obtained in a DRw53 specific assay (which uses a DRw53 fibroblast as the source of Class II molecules). Two very distinct binding patterns were noted, and in several instances, a peptide bound to one DR molecule with high affinity, and did not bind to the other (data not shown).

DR6w19. The DR6w19 assay utilizes as the source of Class II molecules the EBV transformed homozygous cell line H0301, which co-expresses DRB30301 (DR52a). While the radiolabeled ligand used in the DR6w19 assay is different than that used for the DR52a assay, the ligand is related (i.e., is a single substitution analog) to a

high affinity DR52a binder. As was done in the case of DR4w15, the specificity of the assay was investigated by analyzing the binding capacity of a panel of naturally occurring peptides for DR6w19 and DR52a. The two assays demonstrated completely different binding specificities. For example, in terms of relative binding, TT 1272-1284 binds 63-fold better in the DR52a assay than in the DR6w19 assay. Conversely, the Invariant chain peptide binds 189-fold better in the DR6w19 assay. In conclusion, these data demonstrated that the binding of the radiolabeled peptide 650.22 to purified Class II MHC from the H0301 cell line is specific for DR6w19.

DR8w2 and DR8w3. The β1 specificity of the DR8w2 and DR8w3 assays is obvious in that no β3 (and/or B4 and β5) molecule is expressed.

DR9. The specificity of DR9 assay is inferred from previous studies which have shown that the TT 830-843 radiolabeled probe peptide does not bind to DRw53 molecules (Alexander, *et al.*, *Immunity 1:751* (1994)).

Results

DR binding affinity of antigenic peptides recognized by DR restricted T cells

To define a threshold DR binding affinity, to be considered as biologically significant, we compiled the affinities of a panel of 32 reported instances of DR restriction of a given T cell epitope. In approximately half of the cases, DR restriction was associated with affinities of less than 100 nM, and in the other half of the instances, with IC50% in the 100-1000 nM range. Only in 1 out of 32 cases (3.1%) DR restriction was associated with IC50% of 1000 nM or greater. It was noted that this distribution of affinities differs from what was previously reported for HLA class I epitopes, where a vast majority of epitopes bound with IC50% of 50 nM or less (Sette, *et al.*, JI, 1994). This relatively lower affinity of class II restricted epitope interactions might explain why activation of class II restricted T cells in general requires more antigen relative to class I restricted T cells.

In conclusion, this analysis suggested that 1000 nM may be defined as an affinity threshold associated with immunogenicity in the context of DR molecules, and for this reason a suitable target for our studies.

P1 and P6 anchors are necessary but not sufficient for DRB10401 binding

Several independent studies have pointed to a crucial role in DRB10401 binding of a large aromatic or hydrophobic residue in position 1, near the N-terminus of the peptide and of a 9-residue core region (residues 1 through 9). In addition, an important role has been demonstrated for the residue in position six (P6) of this 9-residues core region. Short and/or hydrophobic residues were in general preferred in this position (O'Sullivan, *et al.*, JI 147:2663, 1991; Sette, *et al.*, JI 151:3163, 1993; Hammer, *et al.*, Cell 74:197, 1993 and Marshall, *et al.*, JI 154:5927, 1995).

In the present set of experiments, a library of 384 peptides was analyzed for DRB10401 binding capacity and screened for the presence of the P1-P6 motif (that is, F, W, Y, L, I, V or M in P1 and S, T, C, A, P, V, I, L or M in P6, at least 9 residues apart from the peptide C-terminus. This set of 384 peptides contained a total of 80 DR4w4 binders (specifically 27 good binders [IC50 of 100 nM or less], and 53 intermediate binders [IC50 of the 100-1000 range]. Seventy-seven out of the 80 DR4w4 binders (96%) carried the P1-P6 motif. However, it should be noted that most non-DR4w4 binding peptides also contained the P1-P6 motif. Of 384 peptides included in our database, only 125 were "P1-P6 negative." Only three of them (6%) bound appreciably to purified DR4w4 as opposed to 77/259 (30%) of the "P1-P6 positive" peptides. Therefore, these results demonstrate that presence of suitable P1 and P6 anchors are necessary but not sufficient for DRB10401 binding.

A detailed map of DRB10401 peptide interactions

Next, for each P1-P6 aligned core region, in analogy with what the strategy previously utilized to detail peptide class I interactions the average binding affinity of peptides carrying a particular residue, relative to the remainder of the group, were calculated for each position. Following this method a table of average relative binding (ARB) values was compiled. This table also represents a map of the positive or negative effect of each of the 20 naturally occurring amino acids on DRB10401 binding capacity when occupying a particular position, relative to the main P1-P6 anchors (Figure 1).

Variations in ARB values greater than four fold (ARB ≥ 4 or ≤ 0.25) were arbitrarily considered significant and indicative of secondary effects of a given residue on DR-peptide interactions. Most secondary effects were associated with positions 4, 7, and 9. These positions correspond to secondary anchors engaging shallow pockets on the DR

molecule. In addition, significant secondary effects were detected for M in position 3 (ARB = 12.8) T in position 3 (ARB = 4.34) and I in position 5 (ARB = 4.4).

Development of a DRB10401 specific algorithm

Next, the ARB table was utilized to develop a DRB10401 specific algorithm. In order to predict 0401 binding propensity, each aligned P1-P6 sequence was scored by multiplying, for each position, the ARB value of the appropriate amino acid. According to this procedure, a numerical "algorithm score" was derived. If multiple P1-P6 alignments were possible, binding scores were calculated for each one and the best score was selected. The efficacy of this method in predicting 0401 binding capacity is shown in Table IIa.

Considering only peptides with algorithm scores above -17.00 narrowed the set of predicted peptides to 156. This set still contained 72 out of 80 (90%) of the total high or intermediate DR binders. Raising the cut-off to an algorithm score of -16.44 or higher still allowed identification of 60 out of 80 (75%) of the DR4w4 binding peptides. Of the whole 107 peptide set, twenty-five of them were either good or intermediate binders. In other words, as expected, increasing the algorithm score stringency predicted a smaller fraction of the total binders present in the set, but at the same time less false positive peptides were identified.

Blind test of the predictive power of the DRB10401 specific algorithm

To verify that the predictive capacity of our algorithm was not merely a reflection of having utilized the same data set to test and define the algorithm itself, we further examined its efficacy in a blind prediction test. For this scope we utilized data from an independent set of 50 peptides, whose binding affinities were known, but that had not been utilized in the derivation of the algorithm. As shown in Table IIb, the algorithm was effective in predicting DR4w4 binding capacity of this independent peptide set. The algorithm score of -17.00 identified a total 18 peptides. This set contained 3/3 (100%) of all good binders, and 8/11 (70%) of all intermediate binders in the entire test set of 50 peptides. Increasing the cut-off value to -16.44, identified a set of nine peptides. Seven of them (78%) were either good or intermediate binders. This set contained 7 out of 14

(50%) of the binders contained in the blind prediction peptide set. In conclusion, these data supports the validity of the DR4w4 specific algorithm described above.

Detailed maps of DRB10401, DRB10101, and DRB10701 peptide binding specificities

Next, we analyzed the binding to purified DR1 and DR7 molecules for the same set of 384 peptides utilized to define the DR4w4 algorithm. It was found that this set contained 120 and 59 binders for the DR1 and DR7 alleles, respectively. A total of 158 peptides were capable of binding either DR1, DR4w4 or DR7. A large fraction of them (73/158; 46%) were also degenerate binders, which bound two or more of the three alleles thus far considered. Furthermore, we also found that more than 90% of the DR1 or DR7 good and intermediate binders carried the P1-P6 motif. Most importantly, 72 out of 73 (99%) degenerate DR binders carried this motif (data not shown). In conclusion, this analysis suggests that P1-P6 based algorithms might be utilized to effectively predict degenerate DR binders.

In analogy with what was described above for DR4w4 molecules, specific algorithms were designed for the DR1 and DR7 alleles. Figures 2A and 2B detail the allele specific maps defined according to this method.

As in the case of DRB10401, most secondary effects were concentrated in positions 4, 7 and 9. Position 4 was especially prominent in the case of DR1, while position 7 was the most prominent secondary anchor for DR7. Specific algorithms were developed based on these maps, and it was found that the cut-off values necessary to predict 75% or 90% of the binders were -19.32 and -20.28 for DR1, and 20.91 and -21.63 for DR7, respectively. Depending on the particular allele or cut off value selected, 40 to 60% of the predicted peptides were in fact good or intermediate binders (data not shown).

Development of a DR1-4-7 combined algorithm

Finally, we examined whether a combined algorithm would allow to predict degenerate binders. For this purpose, the sequences of the 384 peptides in our database were simultaneously screened with the three (DR1, 4w4, and 7) specific algorithms. It was found that an even 100 peptides were predicted (using the 75% cut off) to bind either two or three of the alleles considered. This set contained 59 out of 73 (81%) of the

peptides which were in fact capable of degenerate 1-4-7 binding (defined as the capacity to bind to more than one of the DR1, 4w4 or 7 alleles) (Table III).

Definition of a target set of DR specificities, representative of the world population

The data presented in the preceding sections illustrates how peptides capable of binding multiple DR alleles can be identified by the use of a combined "1-4-7" - algorithm. Next, we wished to examine whether the peptides exhibiting degenerate 1-4-7 binding behavior would also bind other common DR types as well. As a first step in our experimental strategy, we sought to define a set of target DR types representative of a large (≥ 80%) fraction of the world population, irrespective of the ethnic population of origin. For this purpose, seven additional DR antigens were considered. For each one of the DR antigens considered in this study, (including DR1, 4 and 7), the estimated frequency in various ethnicities, according to the most recent HLA workshop (11th, 1991) is shown in Table IVa, together with the main subtypes thus far identified.

For the purpose of measuring peptide binding affinity to the various DR molecules, one representative subtype for each DR antigen was chosen (Table I). It should be noted that for most antigens, either one subtype is by far the most abundant, or alternatively a significant degree of similarity in the binding pattern displayed by the different, most abundant subtypes of each DR antigen is likely to exist (see comments column of Table IVb). One exception to this general trend is represented by the DR4 antigen, for which significant differences in peptide specificity between the 0401 and 0405 have been reported. Since both alleles are quite frequent (in Caucasians and Orientals, respectively) we included both DR 0401 and 0405 in the set of representative DR binding assays.

Our set of representative assays is mostly focused on allelic products of the gene, because these molecules appear to be the most abundantly expressed, serve as the dominant restricting element of most human class III responses analyzed thus far, and accurate methods for serologic and DNA typing most readily available. However, we have also considered in our analysis assays representative of DRB3/4/5 molecules (Table IVc). These molecules serve as a functional restriction element, and their peptide binding specificity has been previously shown to have certain similarities to the specificity of several common DRβ_1 allelic products.

A general strategy for prediction of DR-degenerate binders.

To test whether the 1-4-7 combined algorithm would also predict degenerate binding to other common DR types, we measured the capacity of three different groups of synthetic peptides to bind the panel of purified HLA DR molecules. The three different peptide sets were: A) 36 peptides which did not score positive in the combined 1-4-7 algorithm (non-predictions), B) 36 peptides which did score positive for the 1-4-7 algorithm, at the 75% cut off level, but had been found upon actual testing not to be degenerate 1-4-7 binders ("wrong" predictions), and C) 29 peptides which scored positive in the 1-4-7 algorithm, and also proved upon experimental testing, to be actual 1-4-7 degenerate binders (correct predictions). The results of this analysis are shown in Table V.

Within the set of "non-predictions" peptides (Table Va) only 3 out of 34 (9%) bound at least two of the DR1, 4w4 or 7 molecules. Interestingly, 2 (1136.04 and 1136.29) out of 3 of these peptides were also rather crossreactive, and bound additional DR types (DR2w2 β2, DR4w15, 5w11 and 8w2 in the case of 1136.04, and 2w2 β2, 4w15, 9 and 5w12 in the case of 1136.29). Peptides from the "wrong predictions" peptide set (Table V5), by definition bound at the most only one of the DR1, 4w4 or DR7 molecules, and were also poorly degenerate or other DR types with only two peptides (1136.22 and 1188.35) binding a total of three DR molecules. Within this peptide set, no peptide bound four or more of the DR molecules tested (data not shown).

These results are contrasted by data obtained with the peptide set corresponding to peptides which were first predicted by the use of the combined 1, 4, 7 algorithm, and then experimentally found to be degenerate DR1-4-7 binding. Fourteen out of 29 peptides tested (48%) bound a total of five or more alleles. Four of them were remarkably degenerate (1188.16, 1188.32, 1188.34 and F107.09) and bound a total of nine out of the 11 DR molecules tested. In conclusion, these results suggest that a strategy based on the sequential use of a combined DR1, 4, 7 algorithm and quantitative DR1, 4, 7 binding assays can be utilized to identify broadly crossreactive DR binding peptides.

Definition of the HLA-DR 1-4-7 supertype

The data presented above also suggested that several common DR types are characterized by largely overlapping peptide binding repertoires. When this issue was analyzed in more detail, by analyzing the binding pattern of the thirty-two peptides from Table Va and b which were actual DR1-4-7 degenerate binders. Thirty-one of them (97%) bound DR1, 22 (69%) DR4w4 and 21 (66%) DR7. These files are contrasted with the low percentages of binding observed amongst the remainder non-degenerate binding peptides (17/67 (25%), 8/67 (12%) and 7/67 (10%), for DR1, 4w4 and 7, respectively) (Table VII).

Interestingly, a large fraction of the 1-4-7 degenerate binders also bound certain other common DR types. Sixteen (50%) bound DR2w2a, 18 (56%) DR6w19, 18 (56%) DR2w2b and 20 (62%) DR9. In all cases, the frequency of binding in the non-1-4-7 degenerate peptide set was much lower (Table VIII).

Significant, albeit lower, frequencies of cross reactivity were noted also for DR4w15, DR5w11, and DR8w2 (in the 28 to 37% range). Finally, negligible levels of cross reactivity were observed in the case of DR3 and 5w12 and DR53. Further studies will address whether either of these two group of molecules (DR4w15, 5w11, and 8w2 on one hand, and DR3, DR53 and 5w12 on the other) might belong to different DR supertypes.

In conclusion, these data demonstrates that a large set of DR molecules encompassing DR1, 4w4, 2w2a, 2w2b, 7, 9 and 6w19 is characterized by largely overlapping peptide binding repertoires.

Discussion

In the present report we have analyzed the peptide binding specificity of a set of 13 different DR molecules, representative of DR types common among the worldwide population. Detailed maps of secondary anchors and secondary interactions have been derived for three of them (DR4w4, DR1 and DR7). Furthermore, we demonstrated that a set of at least seven different DR types share overlapping peptide binding repertoires; and consequently that broadly degenerate HLA DR binding peptides are a relatively common occurrence. This study also describes computerized procedures which should greatly assist in the task of identification of such degenerate peptides.

We would like to discuss the data in the context of our current understanding of peptide-class II interactions, as well as in the context of the recently described class I supermotifs. Finally, the potential implications of broadly degenerate class II epitopes for epitope based vaccine design should also be considered.

Firstly, our studies illustrate how the vast majority of the peptides binding with good affinity to DR4w4, DR1, DR7 and most of the other DR types analyzed in the current study (data not shown), are all characterized by a P1-P6 motif consistent with the one originally proposed by O'Sullivan, *et al*. Crystallographic analysis of DR1-peptide complexes revealed that the residues occupying these positions engage two complementary pockets on the DR1 molecule, with the P1 position corresponding to the most crucial anchor residue and the deepest hydrophobic pocket. Our analysis also illustrates how other "secondary anchor" positions drastically influence in an allele-specific manner peptide binding capacity. Position 4 was found to be particularly crucial for DR1 binding, position 9 for DR4w4, and position 7 for DR7. These data are consistent with previous results which originally described such allele-specific anchors, and with crystallographic data which illustrates how these residues engage shallow pockets on the DR molecule.

Secondly, our studies illustrate how an approach based on alignment and calculation of average relative binding values of large peptide libraries allows definition of quantitative algorithms to predict binding capacity. The present study extends those observations to two other common HLA-DR types, and also illustrates how the combined use of the 1-4-7 algorithms can be of aid in identifying broadly degenerate DR binding peptides.

The data presented herein suggest that a group of common DR alleles, including at least DR1, DR2w2a, DR2w2b, DR4w4, DR6w19, DR7 and DR9 share a largely overlapping peptide repertoire. Degenerate peptide binding to multiple DR alleles, and recognition of the same epitope in the context of multiple DR types was originally described by Lanzavechia, Sinigallia's and Rothbard's groups. The present study provides a classification of alleles belonging to a main HLA-DR supertype (DR1-4-7-like) which includes DR1, DR2w2a, DR2w2b, DR4w4, DR7, DR9, DR6w19. On the basis of the data presented herein, at least two additional groups of alleles exist. The first group encodes for molecules with significant, albeit much reduced overlap with the 1-4-7-like supertype (DR4w15, 8w2, 5w11). The second group of alleles (5w12, 3w17, and w53)

clearly has little repertoire association with the 1-4-7 supertype. In this context it is interesting to note that Hammer, *et al.* noted that good DR5w11 binding peptides are frequently characterized by positively charged P6 anchor (which would be poorly compatible) with the herein proposed 1-4-7 supermotif. It is also interesting to note that Sidney, *et al.* proposed that DR3w17 binds a set of peptides largely distinct from those bound by other common DR types. Future studies will have to determine whether any of the molecules listed above can be grouped in additional DR supertypes. Our group is currently investigating whether analysis of polymorphic residues lining the peptide binding pockets of DR can be utilized to aid in the classification and prediction of HLA DR supertypes.

We would like to comment on similarities and differences between the HLA DR supertype described herein and the recently described HLA class I supermotifs. Class I supermotifs are clear-cut and, as a rule, non-overlapping. Four of them have been described all approximately equally frequent amongst the worldwide population. By contrast, the repertoire defining the HLA DR supertype herein described is not clear-cut and overlaps, at least in part, with the repertoire of other alleles. It also appears that on the basis of the data presented in Tables I and IV, even if other DR supertypes exist, the DR1-4-7 is going to be by far the most abundantly represented worldwide.

Finally, we would like to point out the possible relevance of these data in terms of development of epitope based vaccines. Class II restricted HTL have been implicated in protection from, and termination of many important diseases. Inclusion of well defined class II epitopes in prophylactic or therapeutic vaccines may allow to focus the immune response towards conserved or subdominant epitopes, and avoid suppressive determinants. Based on the data presented herein (Table IV), the DR1-4-7 supertype would allow coverage in the 50 to 80% range, depending on the ethnicities considered. It is thus possible that broad and not ethnically biased population coverage could be achieved by considering a very limited number of peptide binding specificities.

Based on the results present above, the sequences of various antigens of interest were scanned for the presence of the DR 1-4-7 motifs. Peptides identified using this approach are broadly cross reactive, class II restricted T cell epitopes. Table VIII presents a listing of such peptides derived from HBV, HCV, HIV and *Plasmodium falciparum* (Pf). A total of 146 peptides were identified: 35 from DHBV, 16 from HCV,

50 from HIV, and 45 from Pf. Standard conservancy criteria were employed in applicable cases.

The above examples are provided to illustrate the invention but not to limit its scope. Other variants of the invention will be readily apparent to one of ordinary skill in the art. All publications, patents, and patent applications cited herein are hereby incorporated by reference for all purposes.

Table I

HLA-DR binding assays utilized in the present study.

Antigen	Representative Assay					Comments
	Allele	Alias	Cell Line	Radiolabeled Probe	Ref.	
DR1	DRB1*0101	(DR1)	LG2	HA Y307-319 [1]	(8)	01 is the most prevalent DR1 allele.
DR2	DRB1*1501	(DR2w2b)	L466.1	MBP 88-102Y [2]	(8)	0101 is the most prevalent DR2 allele.
DR3	DRB1*0301	(DR3w17)	MAT	MT 65kD Y3-13 analog [3]	(8)	01 is the most prevalent DR3 allele in most major populations. 01 and 02 are split fairly evenly in NA Blacks.
DR4	DRB1*0401	(DR4w4)	Preiss	Non-natural peptide YAR [4]	(8)	01 is the most prevalent DR4 allele.
	DRB1*0405	(DR4w15)	KT3	Non-natural peptide YAR	This paper	05 is the most prevalent DR4 allele in the Orient.
DR7	DRB1*0701	(DR7)	Pitout	TT 830-843 [5]	(8)	01/02 vary at 1 pos., which is outside the binding groove.
DR8	DRB1*0802	(DR8w2)	OLL	TT 830-843	This paper	02 is dominant in most major population groups. 02 and 03 have nearly identical binding specificities (J. Sidney and A. Sette, unpublished observations).
DR9	DRB1*0901	(DR9)	9074 (HID)	TT 830-843	This paper	DR9 splits are products of a silent mutation.
DR11	DRB1*1101	(DR5w11)	Sweig	TT 830-843	(8)	01 is the most prevalent DR11 allele, by far.
DR12	DRB1*1201	(DR5w12)	Herluf	C1R derived peptide [6]	(9)	01/02 are evenly distributed. These alleles differ at pos. 67, which does not appear strongly influence peptide binding.
DR13	DRB1*1302	(DR6w19)	HO301	650.22 (TT 830-843 analog) [7]	(10)	02 is slightly more prevalent overall than 01. These alleles vary at pos. 86 (critical in determining the P1 anchor specificity.
DR51	DRB5*0101	(DR2w2a)	GM3107	TT 830-843 [5]	(8)	0101 is the most prevalent split.
DR53	DRB4*0101	(DR4, DR7, DR9)	L257.6	Non-natural peptide YAR [4]	(8)	0101 is essentially the only allele.

1) YPKYVKQNTLKLAT
2) VVHFFKNIVTPRTPPY
3) YKTIAFDEEARR
6) EALIHQLKINPYVLS
7) QYIKANAKFIGITE
8) Valli et al., J. Clin. Invest. 91:616, 1993.

Table II

An algorithm to predict DRB1*0401 binding capacity.
a) Original peptide set.

Selection Criteria	No. of peptides (Binding nM) High ≤100	Inter. 100-1000	Non >1000	Total
None	27	53	304	384
P1-P6	27	50	182	259
-17.00 [1)]	27	45	84	156
-16.44 [2)]	25	35	47	107

1) Algorithm score which predicts 90% of all binders.

2) Algorithm score which predicts 75% of all binders.

Table II

b) Blind test of the predictive power of the DRB1*0401 algorithm.

Selection Criteria	No. of peptides (Binding nM) High ≤100	Inter. 100-1000	Non >1000	Total
None	3	11	36	50
P1-P6	3	9	28	40
-17.00	3	8	7	18
-16.44	3	4	2	9

Table III

A combined "1-4-7" algorithm.

Selection Criteria	Degenerate Binders [1]	Percent of Total Degenerate Binders
None	73/384	100%
P1-P6	72/259	99%
Combined Algorithms (90% Cutoff Value)	67/147	92%
Combined Algorithms (75% Cutoff Value)	59/100	81%

1) Degenerate binders are defined as peptides binding at least two out of the three DR1, 4w4, and 7 molecules with an IC50 of 1 μM or less.

Table IV

Phenotypic frequencies of 10 prevalent HLA-DR antigens

Antigen	Alleles	Phenotypic Frequencies					
		Cauc.	Blk.	Jpn.	Chn.	Hisp.	Avg.
DR1	DRB1*0101-03	18.5	8.4	10.7	4.5	10.1	10.4
DR2	DRB1*1501-03	19.9	14.8	30.9	22.0	15.0	20.5
DR3	DRB1*0301-2	17.7	19.5	0.4	7.3	14.4	11.9
DR4	DRB1*0401-12	23.6	6.1	40.4	21.9	29.8	24.4
DR7	DRB1*0701-02	26.2	11.1	1.0	15.0	16.6	14.0
DR8	DRB1*0801-5	5.5	10.9	25.0	10.7	23.3	15.1
DR9	DRB1*09011,09012	3.6	4.7	24.5	19.9	6.7	11.9
DR11	DRB1*1101-05	17.0	18.0	4.9	19.4	18.1	15.5
DR12	DRB1*1201-02	2.8	5.5	13.1	17.6	5.7	8.9
DR13	DRB1*1301-06	21.7	16.5	14.6	12.2	10.5	15.1
Total		97.0	83.9	98.8	95.5	95.6	94.7

Table V

A) Non Predictions.

	Binding Capacity												
	DR1,4,7			Other Alleles									
Peptide	DR1	DR4w4	DR7	DR2w2b	DR2w2a	DR3	DR4w15	DR5w11	DR6w19	DR8w2	DR9	DR5w12	Total Alleles Bound
1136.29	32	4327	138	1.1	468	–	745	6250	–	2970	183	1000	7
1136.04	24	20	3333	1264	741	–	563	69	–	55	2885	–	6
1136.19	781	1915	1323	86	1250	–	445	183	1667	5052	3125	–	4
1136.49	–	–	505	–	702	–	250	645	–	1581	4167	9091	4
1136.02.01a	806	–	–	2844	16	–	–	1379	–	338	–	9927	3
1136.35	116	–	–	2459	–	–	1086	126	8750	306	–	1364	3
1136.52	–	7031	556	3957	1667	–	563	–	–	–	2419	–	2
1136.03	79	8654	2033	243	1250	–	1689	–	–	7313	3947	3571	2
1136.06	1923	1364	–	–	313	6977	–	690	8750	–	–	–	2
1136.23	962	–	–	262	–	2727	–	–	3182	–	–	–	2
1136.32	37	–	–	1717	1739	–	626	6250	–	1976	–	–	2
1136.33	52	–	–	8273	6250	–	7600	1835	8750	3161	–	476	2
1136.44.01	526	780	–	–	–	–	6552	4000	–	6364	–	–	2
1136.62.01a	–	–	–	–	449	–	–	396	–	2970	3000	–	2
1136.42	–	1875	–	–	769	–	–	9524	8750	–	–	–	1
1136.54	8333	–	–	–	–	–	–	–	761	–	–	2727	1
1136.07.01b	1190	–	4630	1542	2857	–	–	1980	–	1225	2614	214	1
1136.05	–	492	–	–	–	–	–	–	–	–	–	–	1
1136.08	–	9375	3788	73	–	–	–	–	–	–	2027	–	1
1136.25	1163	–	6250	28	3846	–	–	–	2917	–	–	3846	1
1136.34	4545	545	3247	–	–	–	–	–	–	–	5000	–	1
1136.36	204	–	–	5688	–	–	–	–	–	–	12931	–	1
1136.64	–	225	–	–	–	–	1267	–	5000	–	–	–	1
1136.69	–	–	–	–	–	–	–	–	54	–	5769	–	1
1136.40	4545	1546	8333	–	4348	–	–	–	7000	–	–	–	0
1136.50	–	1875	–	–	–	–	6667	7143	–	5506	–	–	0
1136.56	–	4500	–	–	–	–	3918	–	3500	–	–	–	0
1136.57	–	8654	–	6500	–	–	5758	1626	–	5104	4688	–	0
1136.61	–	–	–	–	–	–	–	–	–	–	7979	–	0
1136.66	–	–	–	–	–	–	–	–	–	–	–	–	0
1136.68	–	–	–	–	–	–	–	–	–	–	–	–	0
1136.70	–	–	–	–	–	–	–	3704	–	–	–	–	0
1136.72	–	–	–	–	–	–	–	–	–	–	–	–	0
1136.63.01a	–	–	–	–	1905	–	–	7692	–	4298	–	–	0

– indicates binding affinity ≥10,000nM.

2 out of 34 (5.9%) degenerate on 5 or more DR types.

Table V

B) Correct Predictions.

Binding Capacity (IC50% nM)

Peptide	DR1,4,7			Other Alleles									Total Alleles Bound
	DR1	DR4w4	DR7	DR2w2b	DR2w2a	DR3	DR4w15	DR5w11	DR6w19	DR8w2	DR9	DR5w12	
1188.16	3.7	7.1	14	1251	23	–	47	30	428	46	28	–	9
1188.32	3.1	44	167	–	29	–	1402	11	7.1	19	126	851	9
1188.34	14	12	66	370	148	1332	959	2703	3.7	68	19	497	9
F107.09	4.1	14	39	5028	286	–	324	963	469	385	29	–	9
27.412	14	282	138	–	323	–	–	31	20	53	590	2495	8
1188.45	26	9.0	57	260	123	757	1057	2532	3.9	28	16	–	8
1136.16	1.6	214	46	1625	34	–	741	3571	1296	488	68	3409	7
1136.21	2.2	51	52	2844	62	–	270	1212	259	1420	132	–	7
1136.11	0.89	99	9615	603	261	–	84	315	–	529	1974	–	7
27.392	41	449	33	310	2499	–	1668	1203	9.8	883	62	–	6
27.417	56	–	425	210	251	–	–	471	33	2177	859	3243	6
1136.38	70	122	2404	258	741	–	133	4000	1842	–	862	–	6
27.388	50	5737	497	18	1536	–	1410	542	38	–	708	2512	5
27.403	78	4146	207	13	2875	–	–	73	66	1672	423	7321	5
1136.71	5.1	776	96	–	1212	–	950	1538	–	–	375	–	5
1136.14	5.3	4787	100	81	135	–	792	–	1400	–	7732	3488	5
1136.24	182	5844	391	506	9524	–	1357	–	6.5	–	–	–	4
27.384	66	–	281	357	–	–	–	–	65	–	458	–	4
1188.13	116	6923	58	382	–	–	1069	–	0.77	–	142	–	4
F107.10	120	2728	67	807	–	–	1647	–	5.5	–	135	–	4
F107.17	221	388	–	–	–	4878	5705	–	76	7640	299	3478	4
F107.23	163	5713	141	4413	–	–	6770	–	14	–	151	–	4
1136.12	105	720	1429	14	2128	–	1583	–	343	2917	2500	–	4
1136.47	2.2	407	2119	303	755	–	5352	4255	–	–	–	–	3
1136.28	0.23	849	3623	2.2	1481	–	6667	9524	3182	7538	–	4478	3
1136.55	65	138	2451	–	–	–	271	4545	5000	–	–	–	3
1136.59.01a	130	39	–	–	29	–	3140	–	–	–	–	–	2
27.415	2011	754	718	653	–	–	–	6712	2234	8997	–	–	2
1136.46	68	985	5814	–	–	–	–	–	–	–	–	–	

– Indicates binding affinity ≥10,000nM.

16 out of 29 (55%) degenerate on 5 or more DR types.

Table VI

Degenerate "1-4-7" binders.

		Binding Capacity (IC50% nM)												
		DR1,4,7			Other Alleles									
Peptide	Sequence	DR1	DR4w4	DR7	DR2w2b	DR2w2a	DR3	DR4w15	DR5w11	DR6w19	DR8w2	DR9	DR5w12	Total Alleles Bound
1188.34	HNWVNHAVPLAMKLI	+	+	+	+	+	–	+	–	+	+	+	+	10
1188.32	GLAYKFVVPGAATPY	+	+	+	–	+	–	–	+	+	+	+	+	9
1188.16	KSKYKLATSVLAGLL	+	+	+	–	+	–	+	+	+	+	+	–	9
F107.09	KYKLATSVLAGLLGN	+	+	+	–	+	–	+	+	+	+	+	–	9
1188.45	RHNWVNHAVPLAMKL	+	+	+	+	+	+	–	–	+	+	+	–	9
27.412	AYKFVVPGAATPYAG	+	+	+	–	+	–	–	+	+	+	+	–	8
1136.11	VVFPASFFIKLPIILA	+	+	–	+	+	–	+	+	–	+	–	–	7
1136.16	LTSQFFLPALPVFTWL	+	+	+	–	+	–	+	–	–	+	+	–	7
1136.21	IPQEWKPAITVKVLPA	+	+	+	–	+	–	+	–	+	–	+	–	7
1136.29	GPPTALRSFGFAFGYM	+	–	+	+	+	–	+	–	–	–	+	+	7
27.392	SSVFNVVNSSIGLIM	+	+	+	+	–	–	–	–	+	+	+	–	7
27.417	VKNVIGPFMKAVCVE	+	–	+	+	+	–	–	+	+	–	+	–	7
1136.04	LFHYYFLSEKAPGSTV	+	+	–	–	+	–	+	+	–	+	–	–	6
27.388	MRKLAILSVSSFLFV	+	–	+	+	–	–	–	+	+	–	+	–	6
1136.38	SSIIFGAFPSLHSGCC	+	+	–	+	+	–	+	–	–	–	+	–	6
27.403	LVNLLIFHINGKIIK	+	–	+	+	–	–	–	+	+	–	+	–	6
1136.71	EPQGSTYAASSATSVD	+	+	+	–	–	–	+	–	–	–	+	–	5
1136.14	FATCFLIPLTSQFFLP	+	–	+	+	+	–	+	–	–	–	–	–	5
27.384	FNVVNSSIGLIMVLS	+	–	+	+	–	–	–	–	+	–	+	–	5
1188.13	AGLLGNVSTVLLGGV	+	–	+	+	–	–	–	–	+	–	+	–	5
F107.10	LAGLLGNVSTVLLGG	+	–	+	+	–	–	–	–	+	–	+	–	5
1136.47	THHYFVDLIGGAMLSL	+	+	–	+	+	–	–	–	–	–	–	–	4
1136.12	IKLPIILAFATCFLIP	+	+	–	+	–	–	–	–	+	–	–	–	4
F107.23	VFNVVNSSIGLIMVL	+	–	+	–	–	–	–	–	+	–	+	–	4
1136.24	NLSNVLATITTGVLDI	+	–	+	+	–	–	–	–	+	–	–	–	4
F107.17	KFVVPGAATPYAGEP	+	+	–	–	–	–	–	–	+	–	+	–	4
1136.28	LAAIIFLFGPPTALRS	+	+	–	+	–	–	–	–	–	–	–	–	3
1136.55	QEIDPLSYNYIPVNSN	+	+	–	–	–	–	+	–	–	–	–	–	3
1136.59.01a	RVYQEPQVSPPQRAET	+	+	–	–	+	–	–	–	–	–	–	–	3
27.415	NVKYLVIVFLIFFDL	–	+	+	+	–	–	–	–	–	–	–	–	3
1136.46	LWWSTMYLTHHYFVDL	+	+	–	–	–	–	–	–	–	–	–	–	2
1136.44.01	WLFPRFKFVWVTYASW	+	+	–	–	–	–	–	–	–	–	–	–	2
		31	22	21	18	16	1	12	9	18	10	20	3	

+ Indicates binding affinity ≤1000nM.

Table VII

DR Type	Frequency of Binders: 1-4-7 Degenerate Binders (%)	Frequency of Binders: Non 1-4-7 Degenerate Binders (%)
1	31/32 (97)	17/67 (25)
4w4	22/32 (69)	8/67 (12)
7	21/32 (66)	7/67 (10)
9	20/32 (62)	2/67 (3.0)
6w19	18/32 (56)	6/67 (8.9)
2w2ßb	18/32 (56)	16/67 (24)
2w2ßa	16/32 (50)	10/67 (15)
4w15	12/32 (37)	4/67 (6.0)
8w2	10/32 (31)	3/67 (4.5)
5w11	9/32 (28)	6/67 (8.9)
5w12	3/32 (9.4)	4/67 (6.0)
3w17	1/32 (3.1)	0/67 (0)
w53	2/16 (13)	7/43 (16)

Table VIII

Sequence	Source	1st Pos	Conservancy	Predicted 1-4-7
IGPFMKAVCVEVEKT	Pf TRAP	227	100	3
ILSVFFLALFFIIFN	Pf EXP1	3		3
KSKYKLATSVLAGLL	Pf EXP1	71		3
KYKLATSVLAGLLGN	Pf EXP1	73		3
LGNVKYLVIVFLIFF	Pf TRAP	4	100	3
LSVFFLALFFIIFNK	Pf EXP1	4		3
LVNLLIFHINGKIIK	Pf LSA1	13		3
MKILSVFFLALFFII	Pf EXP1	1		3
MRKLAILSVSSFLFV	Pf CSP	2	95	3
NSSIGLIMVLSFLFL	Pf CSP	417	95	3
NVKYLVIVFLIFFDL	Pf TRAP	6	100	3
SFYFILVNLLIFHIN	Pf LSA1	8		3
VFFLALFFIIFNKES	Pf EXP1	6		3
YFILVNLLIFHINGK	Pf LSA1	10		3
YISFYFILVNLLIFH	Pf LSA1	6		3
AGLLGNVSTVLLGGV	Pf EXP1	82		2
ANQLVVILTDGIPDS	Pf TRAP	153	100	2
AYKFVVPGAATPYAG	Pf TRAP	514	80	2
DKELTMSNVKNVSQT	Pf LSA1	81		2
FNVVNSSIGLIMVLS	Pf CSP	413	100	2
FYFILVNLLIFHING	Pf LSA1	9		2
GLAYKFVVPGAATPY	Pf TRAP	512	80	2
GRDVQNNIVDEIKYR	Pf TRAP	25	90	2
HILYISFYFILVNLL	Pf LSA1	3		2
HNWVNHAVPLAMKLI	Pf TRAP	62	80	2
IVFLIFFDLFLVNGR	Pf TRAP	12	100	2
KFVVPGAATPYAGEP	Pf TRAP	516	80	2
KSLLRNLGVSENIFL	Pf LSA1	98		2
KYLVIVFLIFFDLFL	Pf TRAP	8	100	2
LAGLLGNVSTVLLGG	Pf EXP1	81		2
LGNVSTVLLGGVGLV	Pf EXP1	85		2
LIFFDLFLVNGRDVQ	Pf TRAP	15	100	2
LVVILTDGIPDSIQD	Pf TRAP	156	100	2
QLVVILTDGIPDSIQ	Pf TRAP	155	100	2
RGYYIPHQSSLPQDN	Pf LSA1	1669		2
RHNWVNHAVPLAMKL	Pf TRAP	61	80	2
RHPFKIGSSDPADNA	Pf EXP1	107		2
SSVFNVVNSSIGLIM	Pf CSP	410	95	2
VFNVVNSSIGLIMVL	Pf CSP	412	95	2
VKNVIGPFMKAVCVE	Pf TRAP	223	100	2
VKYLVIVFLIFFDLF	Pf TRAP	7	100	2
VSTVLLGGVGLVLYN	Pf EXP1	88		2
WENVKNVIGPFMKAV	Pf TRAP	220	100	2
YKFVVPGAATPYAGE	Pf TRAP	515	80	2

Table VIII

Sequence	Source	1st Pos	Conservancy	Predicted 1-4-7
ENRWQVMIVWQVDRM	HIV1 VIF	2	81	3
ERYLKDQQLLGIWGCS	HIV1 ENV	589		3
ESELVSQIIEQLIKK	HIV1 POL	696	80	3
FRKYTAFTIPSINNE	HIV1 POL	303	93	3
GQMVHQAISPRTLNA	HIV1 GAG	172	88	3
IPEWEFVNTPPLVKL	HIV1 POL	593	93	3
LPPVVAKEIVASCDK	HIV1 POL	770	87	3
NREILKEPVHGVYYD	HIV1 POL	485	87	3
PAIFQSSMTKILEPF	HIV1 POL	336	80	3
PPVVAKEIVASCDKC	HIV1 POL	771	87	3
QEQIGWMTNNPPIPV	HIV1 GAG	276	81	3
QGQMVHQAISPRTLN	HIV1 GAG	171	85	3
SPAIFQSSMTKILEP	HIV1 POL	335	80	3
TLNFPISPIETVPVK	HIV1 POL	176	100	3
VKNWMTETLLVQNAN	HIV1 GAG	348	81	3
VPVWKEATTTLFCAS	HIV1 ENV	54	81	3
WEFVNTPPLVKLWYQ	HIV1 POL	596	93	3
WVKVVEEKAFSPEVI	HIV GAG	187	33	3
YYGVPVWKEATTTLF	HIV1 ENV	51	83	3
ASDFNLPPVVAKEIV	HIV1 POL	765	80	2
ASGYIEAEVIPAETG	HIV1 POL	822	93	2
DFNLPPVVAKEIVAS	HIV1 POL	767	87	2
EAIIRILQQLLFIHF	HIV1 VPR	58	82	2
EKVYLAWVPAHKGIG	HIV1 POL	711	93	2
ETAYFLLKLAGRWPV	HIV POL	838	65	2
EVQLGIPHPAGLKKK	HIV1 POL	268	80	2
FWEVQLGIPHPAGLK	HIV1 POL	266	100	2
GCTLNFPISPIETVP	HIV1 POL	174	100	2
GEIYKRWIILGLNKI	HIV1 GAG	294	85	2
GTVLVGPTPVNIIGR	HIV1 POL	153	100	2
HKAIGTVLVGPTPVN	HIV1 POL	149	93	2
IGTVLVGPTPVNIIG	HIV POL	152	74	2
KRWIILGLNKIVRMY	HIV1 GAG	298	88	2
KVYLAWVPAHKGIGG	HIV POL	712	74	2
LICTTAVPWNASWSNK	HIV1 ENV	607		2
LLQLTVWGIKQLQAR	HIV1 ENV	731	80	2
NFPISPIETVPVKLK	HIV1 POL	178	100	2
PQGWKGSPAIFQSSM	HIV1 POL	329	87	2
PVNIIGRNLLTQIGC	HIV1 POL	161	87	2
QHLLQLTVWGIKQLQ	HIV1 ENV	729	80	2
QQHLLQLTVWGIKQL	HIV1 ENV	728	80	2
SPEVIPMFSALSEGA	HIV1 GAG	197	88	2
TKELQKQITKIQNFR	HIV POL	952	67	2
TVLVGPTPVNIIGRN	HIV1 POL	154	100	2
VEAIIRILQQLLFIH	HIV1 VPR	57	82	2
VIPMFSALSEGATPQ	HIV1 GAG	200	88	2
VNIIGRNLLTQIGCT	HIV1 POL	162	87	2
WGCSGKLICTTAVPWN	HIV1 ENV	601		2
WIILGLNKIVRMYSP	HIV1 GAG	300	88	2
YKRWIILGLNKIVRM	HIV1 GAG	297	88	2
FILVNLLIFHINGKI	Pf LSA1	11		3

Table VIII

Sequence	Source	1st Pos	Conservancy	Predicted 1-4-7
AEDLNLGNLNVSIPW	HBV POL	38	95	3
DLNLGNLNVSIPWTH	HBV POL	40	95	3
GFFLLTRILTIPQSL	HBV ENV	181	80	3
IFLFILLLCLIFLLV	HBV ENV	245	80	3
NLNVSIPWTHKVGNF	HBV POL	45	95	3
PFLLAQFTSAICSVV	HBV POL	523	95	3
RFSWLSLLVPFVQWF	HBV ENV	332	100	3
SPFLLAQFTSAICSV	HBV POL	522	95	3
SVRFSWLSLLVPFVQ	HBV ENV	330	80	3
AFSYMDDVVLGAKSV	HBV POL	546	90	2
AGFFLLTRILTIPQS	HBV ENV	180	80	2
FVQWFVGLSPTVWLS	HBV ENV	342	95	2
GAHLSLRGLPVCAFS	HBV X	50	90	2
GTSFVYVPSALNPAD	HBV POL	774	80	2
GVWIRTPPAYRPPNA	HBV NUC	123	95	2
HLSLRGLPVCAFSSA	HBV X	52	90	2
IIFLFILLLCLIFLL	HBV ENV	244	80	2
ILLLCLIFLLVLLDY	HBV ENV	249	95	2
IVGLLGFAAPFTQCG	HBV POL	636	90	2
KFAVPNLQSLTNLLS	HBV POL	406	95	2
LAQFTSAICSVVRRA	HBV POL	526	95	2
LCLIFLLVLLDYQGM	HBV ENV	252	95	2
LCQVFADATPTGWGL	HBV POL	694	95	2
LHLYSHPIILGFRKI	HBV POL	501	80	2
LLCLIFLLVLLDYQG	HBV ENV	251	95	2
LVLLDYQGMLPVCPL	HBV ENV	258	90	2
LVPFVQWFVGLSPTV	HBV ENV	339	95	2
PLPIHTAELLAACFA	HBV POL	722	80	2
QCGYPALMPLYACIQ	HBV POL	648	95	2
RDLLDTASALYREAL	HBV NUC	28	80	2
SFGVWIRTPPAYRPP	HBV NUC	121	90	2
SVVLSRKYTSFPWLL	HBV POL	750	85	2
VGLLGFAAPFTQCGY	HBV POL	637	95	2
VPNLQSLTNLLSSNL	HBV POL	409	85	2
WPKFAVPNLQSLTNL	HBV POL	404	95	2
KVLVLNPSVAATLGF	HCV	1255	100	3
PTLWARMILMTHFFS	HCV	2870	79	3
ADLMGYIPLVGAPLG	HCV	131	79	2
AVQWMNRLIAFASRG	HCV	1917	100	2
DLELITSCSSNVSVA	HCV	2812	93	2
DLYLVTRHADVIPVR	HCV	1134	79	2
EDLVNLLPAILSPGA	HCV	1882	79	2
FTTLPALSTGLIHLH	HCV	684	79	2
GARLVVLATATPPGS	HCV	1345	79	2
GIQYLAGLSTLPGNP	HCV	1776	100	2
GVNYATGNLPGCSFS	HCV	161	79	2
IQYLAGLSTLPGNPA	HCV	1777	100	2
LHGLSAFSLHSYSPG	HCV	2919	79	2
VNLLPAILSPGALVV	HCV	1885	79	2
VQWMNRLIAFASRGN	HCV	1918	100	2
YKVLVLNPSVAATLG	HCV	1254	100	2

Class II Peptides

Peptide	AA	Sequence	Source
008.00	16	SALLSSDITASVNCAK	HEL 81-96
200.06	16	SALSEGATPQDLNTML	HIV gp25 41-56
213.10	16	NKALELFRKDIAAKYK	Sp. W. myo. 132-147
506.01	20	NKALELFRKDIAAKYKELGY	SW Myo 132-151
506.03	18	ALELFRKDIAAKYKELGY	Sp. W myo. 134-151
506.05	16	ELFRKDIAAKYKELGY	Sp. W myo. 136-151
570.01	16	MAKTIAYDEEARRGLE	Heat Shock Prot
705.06	20	KVYLPRMKMEEKYNLTSVLM	Ova 279-298
717.04	14	YASFVKTTTLRKFT-NH2	combinatorial; DR2 optimized
857.02	20	PHHTALRQAILCWGELMTLA	HBV core 50-69
865.01	15	YKMKMVHAAHAKMKM	OVA KM core extension
F050.03	20	GFYTTGAVRQIFGDYKTTIC	PLP 91-110
F089.01	15	QNILLSNAPLGPQFP	Tyrosinase 56-70
F098.03	20	AAYAAQGYKVLVLNPSVAAT	HCV NS3 1242-1261
F098.04	20	GYKVLVLNPSVAATLGFGAY	HCV NS3 1248-1267
F098.05	14	GYKVLVLNPSVAAT	HCV NS3 1248-1261
F098.06	19	SYVNTNMGLKFRQLLWFHI	HBV Core 87-105
F098.10	12	GLKFRQLLWFHI	HBV Core 94-105
F134.04	20	TLHGPTPLLYRLGAVQNEIT	HCV NS4 1-20
F134.05	20	NFISGIQYLAGLSTLPGNPA	HCV NS4 151-170
F134.08	21	GEGAVQWMNRLIAFASRGNHV	HCV NS4 293-313 (1914-1934)
IA-p5	17	KPVSQMRMATPLLMRPM	Mouse invariant chain 85-101
Tr-28 p1	24	LPKPPKPVSKMRMATPLLMQALPM	Human invariant chain 80-103
27.0279	15	EYLVSFGVWIRTPPA	HBV NUC 117
27.0280	15	GVWIRTPPAYRPPNA	HBV NUC 123
27.0281	15	RHYLHTLWKAGILYK	HBV POL 145
27.0283	15	VPNLQSLTNLLSSNL	HBV POL 409
27.0288	15	WVTVYYGVPVWKEAT	HIV1 ENV 47
27.0293	15	YYGVPVWKEATTTLF	HIV1 ENV 51
27.0294	15	VPVWKEATTTLFCAS	HIV1 ENV 54
27.0295	15	LSGIVQQQNNLLRAI	HIV1 ENV 711
27.0296	15	QQHLLQLTVWGIKQL	HIV1 ENV 728
27.0297	15	QHLLQLTVWGIKQLQ	HIV1 ENV 729
27.0298	15	LLQLTVWGIKQLQAR	HIV1 ENV 731
27.0304	15	QGQMVHQAISPRTLN	HIV1 GAG 171
27.0307	15	SPEVIPMFSALSEGA	HIV1 GAG 197
27.0310	15	QEQIGWMTNNPPIPV	HIV1 GAG 276
27.0311	15	GEIYKRWIILGLNKI	HIV1 GAG 294
27.0312	15	YKRWIILGLNKIVRM	HIV1 GAG 297
27.0313	15	KRWIILGLNKIVRMY	HIV1 GAG 298
27.0314	15	WIILGLNKIVRMYSP	HIV1 GAG 300
27.0315	15	VKNWMTETLLVQNAN	HIV1 GAG 348
27.0322	15	GTVLVGPTPVNIIGR	HIV1 POL 153
27.0324	15	PVNIIGRNLLTQIGC	HIV1 POL 161
27.0326	15	GRNLLTQIGCTLNFP	HIV1 POL 166
27.0328	15	TLNFPISPIETVPVK	HIV1 POL 176
27.0329	15	NFPISPIETVPVKLK	HIV1 POL 178
27.0341	15	FRKYTAFTIPSINNE	HIV1 POL 303
27.0344	15	SPAIFQSSMTKILEP	HIV1 POL 335
27.0345	15	PAIFQSSMTKILEPF	HIV1 POL 336
27.0349	15	QKLVGKLNWASQIYA	HIV1 POL 437
27.0350	15	VGKLNWASQIYAGIK	HIV1 POL 440
27.0351	15	NREILKEPVHGVYYD	HIV1 POL 485
27.0353	15	IPEWEFVNTPPLVKL	HIV1 POL 593
27.0354	15	WEFVNTPPLVKLWYQ	HIV1 POL 596
27.0360	15	EQLIKKEKVYLAWVP	HIV1 POL 705

Class II Peptides

Peptide	AA	Sequence	Source
27.0361	15	EKVYLAWVPAHKGIG	HIV1 POL 711
27.0364	15	HSNWRAMASDFNLPP	HIV1 POL 758
27.0370	15	ASGYIEAEVIPAETG	HIV1 POL 822
27.0372	15	AEHLKTAVQMAVFIH	HIV1 POL 911
27.0373	15	KTAVQMAVFIHNFKR	HIV1 POL 915
27.0377	15	QKQITKIQNFRVYYR	HIV1 POL 956
27.0379	15	KLLWKGEGAVVIQDN	HIV1 POL 982
27.0381	15	ENRWQVMIVWQVDRM	HIV1 VIF 2
27.0382	15	VEAIIRILQQLLFIH	HIV1 VPR 57
27.0384	15	FNVVNSSIGLIMVLS	Pf CSP 413
27.0387	15	MNYYGKQENWYSLKK	Pf CSP 53
27.0388	15	MRKLAILSVSSFLFV	Pf CSP 2
27.0390	15	NSSIGLIMVLSFLFL	Pf CSP 417
27.0392	15	SSVFNVVNSSIGLIM	Pf CSP 410
27.0393	15	MKILSVFFLALFFII	Pf EXP1 1
27.0398	15	FILVNLLIFHINGKI	Pf LSA1 11
27.0400	15	HILYISFYFILVNLL	Pf LSA1 3
27.0402	15	LLIFHINGKIIKNSE	Pf LSA1 16
27.0403	15	LVNLLIFHINGKIIK	Pf LSA1 13
27.0406	15	NLLIFHINGKIIKNS	Pf LSA1 15
27.0408	15	QTNFKSLLRNLGVSE	Pf LSA1 94
27.0412	15	AYKFVVPGAATPYAG	Pf SSP2 514
27.0415	15	NVKYLVIVFLIFFDL	Pf SSP2 6
27.0417	15	VKNVIGPFMKAVCVE	Pf SSP2 223
27.0418	15	WENVKNVIGPFMKAV	Pf SSP2 220
1186.04	15	CSVVRRAFPHCLAFS	HBV POL 534
1186.06	15	FVQWFVGLSPTVWLS	HBV ENV 342
1186.10	15	LAQFTSAICSVVRRA	HBV POL 526
1186.15	15	LVPFVQWFVGLSPTV	HBV ENV 339
1186.18	15	NLSWLSLDVSAAFYH	HBV POL 422
1186.25	15	SFGVWIRTPPAYRPP	HBV NUC 121
1186.26	15	SPFLLAQFTSAICSV	HBV POL 522
1186.27	15	SSNLSWLSLDVSAAF	HBV POL 420
1188.01	15	DKELTMSNVKNVSQT	Pf LSA1 81
1188.13	15	AGLLGNVSTVLLGGV	Pf EXP1 82
1188.16	15	KSKYKLATSVLAGLL	Pf EXP1 71
1188.32	15	GLAYKFVVPGAATPY	Pf SSP2 512
1188.34	15	HNWVNHAVPLAMKLI	Pf SSP2 62
1188.35	15	IGPFMKAVCVEVEKT	Pf SSP2 227
1188.38	15	KYKIAGGIAGGLALL	Pf SSP2 494
1188.45	15	RHNWVNHAVPLAMKL	Pf SSP2 61
F091.15	16	IKQFINMWQEVGKAMY	HIV1 ENV 566
F107.03	15	LQSLTNLLSSNLSWL	HBV POL 412
F107.04	15	PFLLAQFTSAICSVV	HBV POL 523
F107.09	15	KYKLATSVLAGLLGN	Pf EXP1 73
F107.10	15	LAGLLGNVSTVLLGG	Pf EXP1 81
F107.11	15	RHPFKIGSSDPADNA	Pf EXP1 107
F107.14	15	ANQLVVILTDGIPDS	Pf SSP2 153
F107.17	15	KFVVPGAATPYAGEP	Pf SSP2 516
F107.23	15	VFNVVNSSIGLIMVL	Pf CSP 412
35.0093	15	VGPLTVNEKRRLKLI	HBV POL 96
35.0096	15	ESRLVVDFSQFSRGN	HBV POL 387
35.0100	15	LCQVFADATPTGWGL	HBV POL 683
35.0106	15	VVVVATDALMTGYTG	HCV 1437
35.0107	15	TVDFSLDPTFTIETT	HCV 1466
35.0125	15	AETFYVDGAANRETK	HIV POL 619

Class II Peptides

Peptide	AA	Sequence	Source
35.0127	15	EVNIVTDSQYALGII	HIV POL 674
35.0131	15	WAGIKQEFGIPYNPQ	HIV POL 874
35.0133	15	GAVVIQDNSDIKVVP	HIV POL 989
35.0135	15	YRKILRQRKIDRLID	HIV VPU 31
35.0171	15	PDSIQDSLKESRKLN	Pf SSP2 165
35.0172	15	KCNLYADSAWENVKN	Pf SSP2 211
1280.02	15	IGTVLVGPTPVNIIG	HIV POL 152
1280.03	15	KVYLAWVPAHKGIGG	HIV POL 712
1280.04	15	TKELQKQITKIQNFR	HIV POL 952
1280.06	15	AGFFLLTRILTIPQS	HBV ENV 180
1280.08	15	GFFLLTRILTIPQSL	HBV ENV 181
1280.09	15	GTSFVYVPSALNPAD	HBV POL 774
1280.12	15	IIFLFILLLCLIFLL	HBV ENV 244
1280.13	15	KFAVPNLQSLTNLLS	HBV POL 406
1280.15	15	LHLYSHPIILGFRKI	HBV POL 501
1280.16	15	LLCLIFLLVLLDYQG	HBV ENV 251
1280.21	15	VGLLGFAAPFTQCGY	HBV POL 637
1280.22	15	FYFILVNLLIFHING	Pf LSA1 9
1280.23	15	KSLLRNLGVSENIFL	Pf LSA1 98
1280.25	15	RGYYIPHQSSLPQDN	Pf LSA1 1669
1283.02	15	VYLLPRRGPRLGVRA	HCV Core 34
1283.10	15	GHRMAWDMMMNWSPT	HCV E1 315
1283.11	15	CGPVYCFTPSPVVVG	HCV NS1/E2 506
1283.12	15	VYCFTPSPVVVGTTD	HCV NS1/E2 509
1283.13	15	GNWFGCTWMNSTGFT	HCV NS1/E2 550
1283.14	15	FTTLPALSTGLIHLH	HCV NS1/E2 684
1283.16	15	SKGWRLLAPITAYAQ	HCV NS3 1025
1283.17	15	DLYLVTRHADVIPVR	HCV NS3 1134
1283.20	15	AQGYKVLVLNPSVAA	HCV NS3 1251
1283.21	15	GYKVLVLNPSVAATL	HCV NS3 1253
1283.22	15	VLVLNPSVAATLGFG	HCV NS3 1256
1283.24	15	GARLVVLATATPPGS	HCV NS3 1345
1283.26	15	DVVVVATDALMTGYT	HCV NS3 1436
1283.30	15	FTGLTHIDAHFLSQT	HCV NS3 1567
1283.31	15	YLVAYQATVCARAQA	HCV NS3 1591
1283.33	15	LEVVTSTWVLVGGVL	HCV NS4 1658
1283.34	15	TWVLVGGVLAALAAY	HCV NS4 1664
1283.36	15	AKHMWNFISGIQYLA	HCV NS4 1767
1283.37	15	IQYLAGLSTLPGNPA	HCV NS4 1777
1283.44	15	MNRLIAFASRGNHVS	HCV NS4 1921
1283.50	15	SYTWTGALITPCAAE	HCV NS5 2456
1283.55	15	GSSYGFQYSPGQRVE	HCV NS5 2641
1283.57	15	LELITSCSSNVSVAH	HCV NS5 2813
1283.61	15	ASCLRKLGVPPLRVW	HCV NS5 2939
1298.02	15	VGNFTGLYSSTVPVF	HBV POL 53
1298.03	15	TNFLLSLGIHLNPNK	HBV POL 568
1298.04	15	KQCFRKLPVNRPIDW	HBV POL 615
1298.06	15	KQAFTFSPTYKAFLC	HBV POL 661
1298.07	15	AANWILRGTSFVYVP	HBV POL 764
1298.08	15	PDRVHFASPLHVAWR	HBV POL 824
1298.10	15	IRPVVSTQLLLNGSL	HIV1 ENV 333
1298.11	15	RSELYKYKVVKIEPL	HIV1 ENV 637
1298.13	15	DRFYKTLRAEQASQE	HIV1 GAG 333
1298.16	15	KVILVAVHVASGYIE	HIV1 POL 813
F125.02	17	LVNLLIFHINGKIIKNS	Pf LSA1 13
F125.04	16	RHNWVNHAVPLAMKLI	Pf SSP2 61

WHAT IS CLAIMED IS:

1. A composition comprising an isolated peptide that induces a CTL response and a T helper peptide comprising a motif of about nine residues wherein the first position from the N terminus of the motif is Y, F, W, L, I, V, M and the sixth position from the N terminus of the motif is S, T, C, A, P, V, I, L, M.

2. The composition of claim 1, wherein the T helper peptide consists of between about 10 and about 24 residues.

3. The composition of claim 1, wherein the T helper peptide is derived from a viral antigen.

4. The composition of claim 3, wherein the viral antigen is from HIV, HBV, or HCV.

5. The composition of claim 1, wherein the T helper peptide is derived from a parasite.

6. The composition of claim 5, wherein the antigen is *Plasmodium falciparum*.

7. The composition of claim 1, wherein the peptide that induces a CTL response is linked to the T helper peptide.

8. A method of inducing a CTL response in a patient, the method comprising contacting a cytotoxic T cell from the patient with an isolated peptide that induces a CTL response and a T helper peptide comprising a motif of about nine residues wherein the first position from the N terminus of the motif is Y, F, W, L, I, V, M and the sixth position from the N terminus of the motif is S, T, C, A, P, V, I, L, M.

9. The method of claim 8, wherein the step of contacting is carried out by administering to the patient a pharmaceutical composition comprising the nucleic acid encoding the peptide that induces a CTL response and the T helper peptide.

10. The method of claim 8, wherein the the peptide that induces a CTL response is linked to the T helper peptide.

11. A composition comprising a peptide as shown in Table VIII.

12. A method of inducing a helper T cell response in a patient, the method comprising contacting a helper T cell with a peptide of claim 11.

13. The method of claim 12, wherein the step of contacting is carried out by administering to the patient a pharmaceutical composition comprising the peptide.

14. The method of claim 12, wherein the step of contacting is carried out by administering to the patient a pharmaceutical composition comprising a nucleic acid encoding the peptide.

Figure I

DR4w4 Algorithm: Average Relative Binding Values.

Residue	p1 Anchor	2	3	4	5	p6 Anchor	7	8	9
C		0.57	0.74	1.12	0.83	0.47	0.94	0.28	1.10
G		1.14	0.64	0.43	0.48		0.49	1.19	0.52
S		1.55	1.31	1.29	1.76	1.11	1.23	2.93	1.54
T		1.00	4.34	0.89	1.32	1.86	3.07	1.76	1.64
P		0.56	0.31	1.44	2.46	0.86	2.83	2.12	2.18
A		0.96	1.04	1.57	0.59	0.65	0.86	0.82	1.62
L	0.81	0.86	1.88	1.28	1.11	0.67	1.36	1.08	0.83
I	0.79	1.74	1.01	1.91	4.39	0.98	2.36	1.66	2.75
V	0.79	3.34	0.93	1.05	0.70	2.36	0.69	0.54	1.53
M	1.14	12.79	1.49	2.77	0.32	0.74	8.11	1.98	4.05
F	2.33	3.66	1.85	0.80	1.58		1.84	1.34	1.12
W	0.82	2.04	2.52	0.21	0.91		0.39	0.35	0.21
Y	1.07	0.74	1.51	0.39	1.41		0.44	0.61	0.35
H		0.78	0.15	1.14	0.93		13.77	1.40	5.15
R		1.09	0.50	0.69	0.39		0.14	0.41	1.22
K		1.44	1.25'	0.53	0.40		0.62	0.64	0.55
Q		0.40	0.38	1.61	2.09		0.31	0.71	0.62
N		0.44	1.72	1.42	1.89		0.84	0.43	1.64
D		0.34	0.33	1.40	0.40		0.58	0.53	0.24
E		0.31	1.09	0.42	0.42		0.29	0.61	0.25

Figure II

a) DR1 Algorithm: Average Relative Binding Values.

Residue	p1 Anchor	2	3	4	5	p6 Anchor	7	8	9
C		0.22	0.15	0.49	0.06	0.14	0.31	0.45	0.35
G		1.29	3.38	2.13	1.73		0.23	1.58	0.44
S		0.87	0.48	0.32	0.58	0.74	1.03	1.25	1.03
T		0.57	2.08	0.30	1.59	1.26	1.51	1.73	2.32
P		0.43	0.88	5.42	2.57	0.63	1.78	1.63	1.52
A		1.93	3.51	4.14	1.59	2.42	1.89	1.25	4.09
L	0.97	1.20	0.64	3.08	2.32	0.85	2.02	3.10	0.83
I	1.00	3.84	1.59	1.10	1.30	0.75	3.47	0.67	1.32
V	0.74	2.95	1.08	0.79	1.97	1.16	2.89	0.57	5.89
M	2.82	1.07	2.62	7.66	0.93	2.67	7.27	1.01	4.39
F	1.51	2.05	0.49	0.22	0.40		0.91	0.89	0.79
W	0.30	0.63	0.69	0.56	0.14		0.61	0.35	0.58
Y	0.88	0.51	1.22	0.36	2.04		0.99	0.26	0.42
H		0.51	0.11	0.68	1.57		1.81	1.20	0.55
R		0.80	0.49	0.43	0.37		1.08	1.43	0.83
K		2.69	2.32	0.49	0.67		1.33	2.24	0.44
Q		1.38	1.27	7.07	1.58		1.06	3.65	1.54
N		0.63	1.41	1.20	0.75		1.16	0.43	1.15
D		0.85	0.31	0.20	0.21		0.11	0.08	0.39
E		0.31	0.47	0.59	0.57		0.16	0.53	0.27

b) DR7 Algorithm: Average Relative Binding Values.

Residue	p1 Anchor	2	3	4	5	p6 Anchor	7	8	9
C		0.17	0.58	0.30	0.26	0.45	1.38	0.53	1.04
G		0.45	0.43	0.25	0.54		0.23	1.30	0.22
S		1.86	0.66	1.11	2.39	1.14	1.95	1.67	0.89
T		0.72	6.53	1.88	1.78	0.79	1.54	0.94	1.92
P		0.36	0.37	2.01	0.46	0.49	1.06	0.60	1.78
A		1.43	2.63	4.78	0.89	1.51	0.74	0.89	0.61
L	0.87	1.04	1.08	1.09	0.83	0.89	1.88	1.18	0.97
I	0.77	1.99	0.96	2.17	2.88	1.11	1.11	1.52	5.69
V	0.82	2.15	0.47	0.57	0.92	2.25	1.36	0.80	5.49
M	1.45	5.75	2.54	3.74	0.33	1.21	9.03	3.01	3.42
F	1.97	1.43	0.68	0.90	1.07		2.50	2.39	1.90
W	0.93	1.32	4.07	0.81	0.58		0.81	0.95	0.66
Y	0.90	0.78	3.34	0.62	3.32		0.64	0.74	0.74
H		1.67	0.36	0.62	2.09		1.10	1.02	1.13
R		1.29	0.70	0.45	1.31		0.21	0.59	2.67
K		1.45	1.32	0.47	0.86		1.40	1.26	0.48
Q		1.70	0.82	2.09	1.40		1.01	2.68	0.36
N		1.42	2.35	0.86	1.68		1.62	0.24	0.88
D		0.61	0.41	0.27	0.26		0.19	0.44	0.30
E		0.48	0.59	1.23	0.74		0.45	0.57	1.16

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/01373

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :424/185.1; 514/14,13, 44; 535/23.4

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 424/185.1; 514/14,13, 44; 535/23.4

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE, SCISEARCH, EMBASE, BIOSIS, MHC, binding motif, DR1, Dr4, Dr7

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO94/03205 A1 (CYTEL CORPORATION) 17 February 1994, see entire document.	1-14
Y	WO 93/20103 A2 (ISIS INNOVATION LIMITED) 14 October 1993, see entire document.	1-14
Y	WO 95/26982 A1 (ISIS INNOVATION LIMITED) 12 October 1995, see entire document.	1-14
Y	WO 95/25122 A1 (THE SCRIPPS RESEARCH INSTITUTE) 21 September 1995, see entire document.	1-14

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
01 MAY 1998	01 JUN 1998

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
MARTHA LUBET
Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)★

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/01373

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
first species

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

A61K 38/08, 38/10, 39/10,39/02, 39/12; C07K 7/00, 14/005, 14/20, 14/195, 14/725; C07H 21/04

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for more than one species to be searched, the appropriate additional search fees must be paid. The species are as follows: Each of the peptides claimed in claim 11 (the peptides listed in Table VIII).

The species listed above do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, the species lack the same or corresponding special technical features for the following reasons: Each of the peptides listed in Table VIII have unique amino acid sequences. Each of the peptides of Table VIII have distinct biochemical structure and properties, ie ability to bind to a particular HLA molecule or elicit peptide specific antibody. Accordingly, the peptides disclosed in Table VIII are not so linked by a special technical feature within the meaning of PCT Rule 13.2 so as to form a general inventive concept.